Empathy : the cognitive and neural correlates

Lawrence, Emma Jane

Empathy: Cognitive and Neural Correlates

Emma J Lawrence

Department of Psychological Medicine
Institute of Psychiatry, King's College
University of London

Thesis submitted for the degree of Doctor of Philosophy

2004

Acknowledgements

Firstly, I would like to especially thank those with Asperger's Syndrome and those who suffer with depersonalisation who volunteered their time to take part in the studies detailed in this thesis, without whom none of this would be possible.

I am extremely grateful to my supervisor, Professor Tony David, for the invaluable support, advice and guidance he provided at just the right times.

Thanks go to everyone in the Section of Cognitive Neuropsychiatry who provided a supportive, stimulating and enjoyable working environment especially to Philip and Maxine (extra thanks for being so generous with your time and ideas), Dawn and Clare; all of whom have helped in different ways. I am also very grateful those who proofread especially Richard who shared his considerable expertise.

I am indebted to Professor Declan Murphy and the rest of the Section of Brain Maturation for their help and expertise regarding Asperger's Syndrome. And to all who were charitable with their time and resources in helping me master fMRI, particularly Vincent, Simon, Nick and Tamara who helped with the analysis and the radiographers including Carl Shew who has sadly passed away.

Special thanks goes to Alex who went through the entire process with me, encouraging me along the way, to Sarah and Stacey for being there for me as always, and to the rest of the gang who kept me smiling.

Abstract

The cognitive and neural bases of empathy were investigated by collecting data from healthy individuals, people with Asperger's Syndrome (AS) and those with Depersonalisation Disorder (DPD). A principal component analysis was conducted on a self-report measure, the Empathy Quotient (Baron-Cohen & Wheelwright, in press) and the emerging factors were consistent with those identified by social psychologists: cognitive empathy, emotional reactivity and social skills. In chapter 3, a series of tests confirmed that neither clinical group was impaired in reasoning outside the affective domain. In chapters 4 and 5, cognitive tasks were used to tap the different facets of emotion attribution. A speech rate task was developed to implicitly measure emotional arousal and a self-other trait overlap task to gauge the role of the self-concept. Speech rate showed the DPD group did not experience emotional arousal congruent with the target, unlike controls. Both clinical groups displayed more self-other overlap than controls on the trait-overlap task, indicative of a self-orientated strategy. In chapter 6, 'theory of mind tasks' were used tapping the attribution of false beliefs from vignettes (Rowe *et al*, 2001) and complex mental states from photographs of eyes (Baron-Cohen *et al*, 2001). These data suggested the different components of empathy may be 'dissociated'. People suffering DPD showed a disruption in affective empathy with preservation of the cognitive aspects whilst those with AS displayed the opposite profile. Lastly, a novel functional magnetic resonance imaging paradigm was used wherein healthy volunteers inferred mental states from brief video clips of an actor in various scenarios. This was designed to uncover the neural substrates of 'shared representations' or 'simulation'. Brain activation was consistent with action perception models including pre-motor and parietal areas. Overall, this thesis has helped to refine and validate the concept of empathy including the neural correlates.

Table of Contents

List of Tables

List of Graphs & Figures

Abbreviations

AN(C)OVA	Analysis of (co)variance
AS	Asperger's syndrome
ASD	Autistic spectrum disorder
BA	Brodmann area
BAI	Beck anxiety inventory
BDI	Beck depression inventory
DES	Dissociative experiences scale
DLPFC	Dorso-lateral prefrontal cortex
DPD	Depersonalisation disorder
EMB	Extreme male brain hypothesis
EMG	Electromyography
EQ	Empathy quotient
fMRI	Functional magnetic resonance imaging
FPQ	Fundamental power quotient
GSR	Galvanic skin response
HFA	High-functioning autism
IAPS	International affective picture system
IOP	Institute of Psychiatry
IRI	Interpersonal reactivity index
MEG	magnetoencephalography
MS	Mental state
NART	National adult reading test
PA	Propositional attitude
PCA	Principal components analysis
PET	Positron emission tomography
PONS-r	Profile of non-verbal sensitivity – revised
QMEE	Questionnaire measure of emotional empathy
SAS	Self absorption scale
SD	Standard deviation
SDS	Social desirability scale
SQ	Systemising quotient
ToM	Theory of mind

Introduction

Contents at a glance

Empathy: definitions and concepts

We all have an idea of what empathy is, and appreciate its importance. But empathy is a slippery concept that scientists and laypeople alike do not fully understand. Empathy is a multi-dimensional construct (Davis, 1980) with two main components: i) cognitive empathy, 'the intellectual/imaginative apprehension of another's mental state (Hogan, 1969) and ii) emotional empathy, 'a vicarious emotional response to the perceived emotional responses of others' (Mehrabian and Epstein, 1972). The latter can also occur without a conscious representation known as 'emotional contagion' (Doherty, 1998). These definitions have recently been reformulated and updated as: i) cognitive empathy, '*understanding and predicting* someone else's mental state' and ii) affective empathy, '*experiencing an emotion* as the result of someone else's mental state ~~or more broadly~~ (Baron-Cohen and Wheelwright, 2004 ~~in press~~), that emotion having to be appropriate (see below).

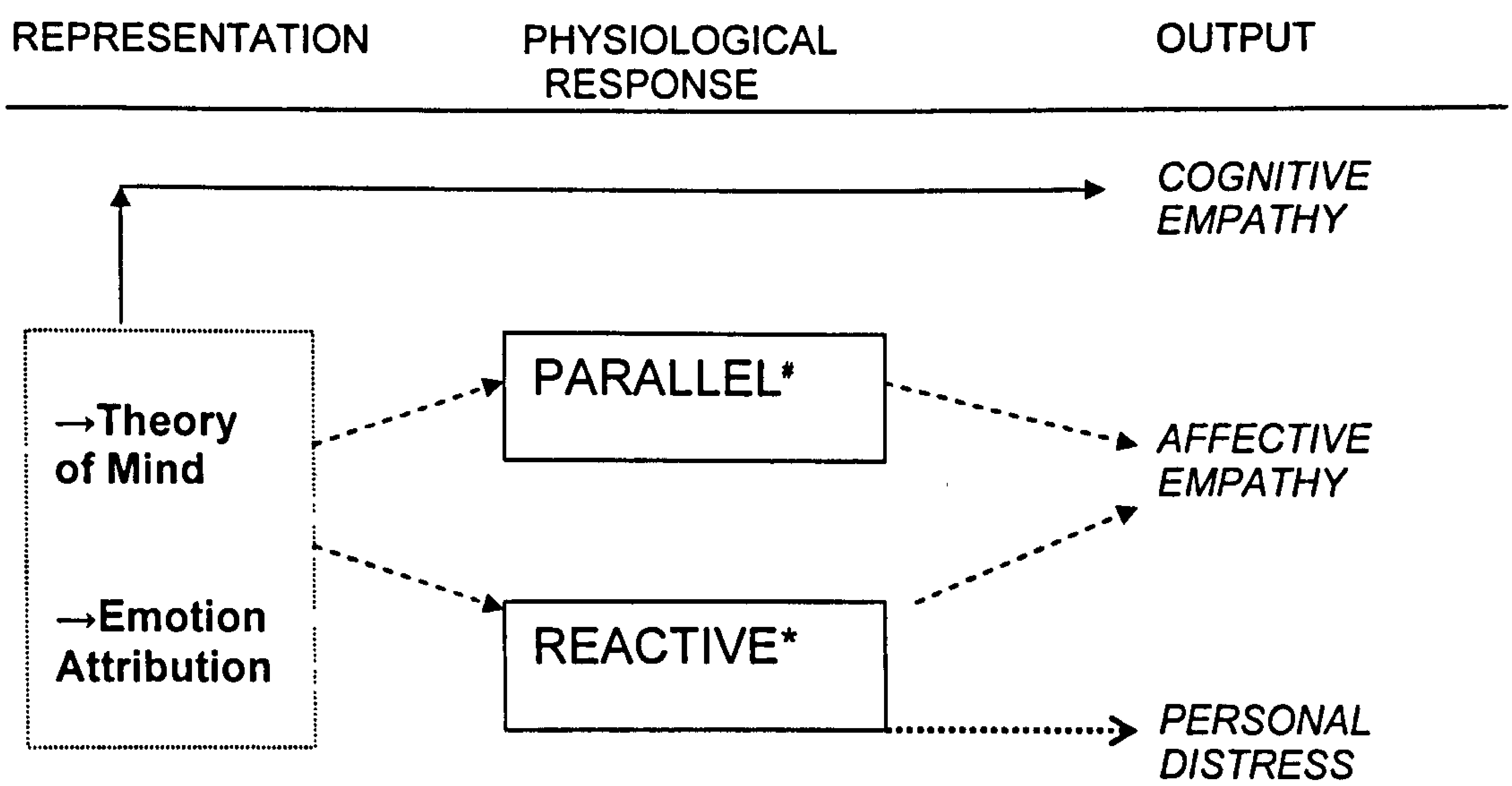
The word 'empathy' is the English translation by Titchener in 1909 of the German word 'Einfühlung' (Baron-Cohen and Wheelwright, 2004 ~~in press~~). 'Einfühlung' is a concept borrowed from aesthetics meaning 'to project yourself into what you observe'. The concept was extended to the study of inter-subjectivity by Theodore Lipps, who conceptualised it in terms of a kind of 'inner imitation' (for an overview see Gallese, 2003). The term sympathy was in use much earlier, and was often used to denote what we now call empathy (Darwin, 1872). Nowadays, academics usually distinguish empathy from sympathy[2] – the latter being defined as 'recognising another's difficulties and being motivated to alleviate them' (Goldie, 1999). However, it is not clear to what degree this distinction is made in ordinary usage.

Titchener was a psychologist concerned with describing and understanding empathy. Over the years, psychologists from many different fields have researched and studied empathy. Personality theorists are interested in empathy as a trait, including how it relates to concepts such as emotional intelligence (Goleman, 1995), and in the relationship ~~of~~ between trait vs. state empathy (Nezlek *et al.*, 2001). Social psychologists have examined its role in group processes and in relation to pro-social behaviour (Davis *et al.*, 1999). Developmental psychologists, on the other hand, see it as an important milestone in areas such as moral reasoning (Harris, 1989). They are also concerned with the link between the development of empathy, parenting and attachment patterns (Van der Mark *et al.*, 2002). Psychotherapists and counsellors take a more experiential interest in empathy as an important part of the therapeutic process (Rogers, 1992). Clinical psychologists and psychiatrists, however, are more concerned with empathy (or its disruption) as a symptom. More recently, neuroscientists have begun to consider the neural substrates of empathy and the effects on this capacity of brain damage or modulation. To this end, it is necessary to further disentangle the concept of empathy.

[2] Although this is not necessarily the case in non English speaking countries – see Decety (2003)

A key requirement for 'affective empathy' is that the emotional response be *appropriate*. The empathic emotional response can be classified in two ways (Davis, 1994): i) 'parallel empathy' the response matches that of the target: for instance, feeling fear at another's fright, and ii) 'reactive empathy' which involves going beyond a simple matching of affect such as sympathy or compassion (see Figure 1.1). However, some emotional responses are not empathic e.g. happiness at another's misfortune or less obviously 'personal distress' (Davis, 1980, Eisenberg and Strayer, 1987).

Figure 1.1: Empathy and related processes



+ 'Theory of mind' is used in the restricted sense to refer to meta-representation (see below).
* feeling sympathy, compassion, pride etc.
" feeling happy in response to another's happiness or fright at another's fear. i.e. direct correspondence with target mental state.

'Personal distress' occurs when a person perceives the misfortune of another and has a self-orientated response stemming primarily from thoughts as to how they would feel in the same situation (Batson *et al.*, 1987) as opposed to an other-orientated response such as sadness, sympathy or concern. Although this is still a type of emotional reactivity, and may share some of the same mechanisms as empathy, it is clearly not empathy itself. More contemporary definitions of affective empathy incorporate this distinction (Hoffman, 2000).

Affective empathy relates purely to emotional states[3], but cognitive empathy encompasses all classes of mental state such as beliefs, desires, emotions and perhaps sensations[4]. A belief is a term for the mental state (MS) of holding a proposition or an idea as true (Blackburn, 1996). They are often referred to as a epistemic or intentional states as they refer to something in the world[5] (Stone *et al.*, 2003). Beliefs are also known as 'propositional attitudes' (PA) because understanding them is a simple case of knowing what they are about or directed towards. To interpret behaviour as governed by intentional states is known as taking the intentional stance (see Griffin and Baron-Cohen, 2002 for discussion). 'Affective states' such as desires and emotions (Stone *et al.*, 2003) also refer to something in the world (and so have propositional content), but have an important qualitative dimension too. This is especially true of emotions, which unlike beliefs cannot be fully appreciated by simply knowing what they are directed towards.

The object of cognitive empathy can be any one of the mental states just described. One can have an 'intellectual or imaginative' understanding of another's belief, desire or an emotion. This may mean that cognitive empathy is in itself a multi-dimensional

[3] Although it is conceivable that someone could have an emotional response to another's belief.
[4] Percepts are also a class of mental state, but they will not be discussed here.
[5] This is also known as having 'content' or intentionality Brentano (1895) *Psychology from an empirical standpoint*, Routledge, London.

construct. Whether all types of mental state attribution share the same cognitive processes and neural structures is currently under debate (Stone *et al.*, 2003). It remains unclear whether 'cognitive empathy' draws on different resources and processes as a function of the target mental state, although there is some evidence that it is not made up of completely disparate abilities (Stone *et al.*, 2003).

The umbrella term 'theory of mind' (ToM) refers to the ability to attribute mental states to others in order to explain and predict behaviour. This definition encapsulates cognitive empathy and the representational component of 'affective empathy'[6] (Blair, 1996). ToM involves both 'cold cognition' such as inferences about epistemic states like beliefs and 'hot cognition' which focuses on affective states such as emotions (Stone, 1999). There is much research on cold ToM but less on how a 'theory of mind' deals with affective states. This is in part due to the difference between emotions and propositional attitudes (PA).

A vital function of a ToM is the ability to appreciate the *representational* qualities of mental states. A fully-fledged ToM must be able to detect that some mental states refer to *representations* of the real world rather than the real world itself. This is especially true of PA: 'I believe it is raining', is true irrespective of the actual weather. For this reason PA are known as *opaque*. Conversely, sensations and perceptions are transparent inasmuch as they refer to the real world and do not ordinarily require a suspension of truth: 'it is raining', is either true or not. Some restrict the use of the term ToM to refer to instances that require the use of meta-representation.

[6] Emotional contagion may be the exception.

Emotions also have a representational component, for example, appreciating that a person is scared of injections requires an understanding of their specific representation of an injection, which may feasibly differ from the injection's 'real-world' qualities. However, in some situations certain aspects of the phenomenal qualities of emotions can be *transparent* inasmuch as facial expressions which explicitly convey information such as valence and intensity (and so automatically give the information regarding the mental state – see Zahavi and Parnas, 2003) are universal (Ekman, 1999). Emotions are therefore a less uniform class of mental state. Epistemic states, however, and in particular false beliefs, provide the perfect litmus test for the operation of a fully competent ToM. They are the only class of mental state which unambiguously tap the ability to employ a meta-representational ToM. This has led to a large amount of data on the comprehension of other peoples' beliefs within this field (Baron-Cohen, 2000).

Our knowledge of emotional understanding has, however, been largely based on simple emotion attribution paradigms using emotional expressions alone without reference to the emotions' content and hence not requiring meta-representation. In terms of empathy, it is not clear where simple emotion expression paradigms fit in. Goldie (1999) argues that simply knowing someone is scared is not enough to be considered an 'understanding' (as in the definition of cognitive empathy), but there has been a recent move to view such paradigms as tapping empathy [see Carr *et al*, (2003) for instance]. Furthermore, it is possible that emotional resonance occurs in such paradigms, suggesting they may also invoke 'affective empathy'. Phillips *et al.*, (2003) point out that producing and regulating an emotional response to stimuli are important parts of emotion perception; however, the issue of emotion induction in emotion perception paradigms is rarely considered[7].

This split between emotions and other mental states has led to a gap in our knowledge of emotional understanding. Ecologically valid cognitive tests which tap the appreciation of affective states in more complex situations other than simple emotional expression paradigms (with or without emotional resonance) have only recently gained popularity (Channon *et al.*, 2001, Harris, 1989, Stone *et al.*, 1998). As a result we do not fully understand the mechanisms involved in such tasks nor whether (or to what degree) they can be dissociated from those used to comprehend epistemic states. Yet it is the affective aspects of empathy that most people think of as defining empathy.

2. Empathy: mechanisms and processes

The definitions and explanations given so far say little about the mechanisms and processes that enable an empathic reaction. Inter-subjectivity may be dependent on a whole range of skills, and the mechanisms involved may depend on the target mental state, the available cues and individual differences. The nature of mental state attribution as a whole has been the focus of debate in the philosophical literature. There are two main schools of thought; however, much of the debate as to the utility of these approaches has focused on propositional attitudes or epistemic states as opposed to affective states (see above).

[7] But see Adolphs, R., Damasio, H., Tranel, D., Cooper, G. and Damasio, A. R. (2000) A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping *Journal of Neuroscience*, **20**, 2683-2690. Also Blair, J. (1999) Psychophysiologial responsiveness to the distress of others in children with autism *Personality & Individual Differences*, **26**, 477- 485.

Theory-theory: this school of thought suggests that people possess or develop a 'theory' to explain others' mental states and behaviours 'using generalisations about how people reason, decide and act' (Henderson, 1996) The 'theory-theory approach'[8] has been described as a 'detached ... activity' (Gallese and Goldman, 1998). This view has been popular amongst functionalists and developmental psychologists. According to Gopnik and Goldman (1993) ascriptions of mental states are regarded as 'theoretical' merely because such states are not directly observable in others[9], and "are seen as 'theoretical posits' invoked to explain and predict behaviour in the same fashion that physicists appeal to electrons and quarks" (Gallese and Goldman, 1998). However, other theory-theorists hold that the theory is innate and arises from a dedicated module as opposed to being constructed from incoming data like as scientific theory (Zahavi and Parnas, 2003).

Simulation: the alternative approach suggests that 'attributors use their own mental mechanisms to calculate and predict the mental processes of others' (Gallese and Goldman, 1998). Gallese *et al.*, (1998) point out that simulation is 'an attempt to replicate, mimic or impersonate the mental life of the target'. When a person is simulating there is said to be a 'correspondence' with the MS of the target with the main difference being that the attributor is using 'pretend' or 'offline' mental states. Imitation or 'overt action simulation' (Blakemore and Decety, 2001) is the likely candidate for the developmental basis for simulation and perhaps ToM *per se* (Meltzoff and Gopnik, 1993b).

[8] Some authors reserve the term 'theory of mind' to refer to the 'theory-theory' approach. However, many use the phrase more broadly to denote any mental state attribution as is the case here.
[9] This is holds true for some aspects of emotions.

Recent formulations suggest that neither account of mental state attribution is likely to be mutually exclusive (Davies and Stone, 1995, Vogeley *et al.*, 2001). It is important to appreciate that both approaches can conceivably be used to understand affective states – with or without emotional resonance The full extent of this debate is outside the scope of this thesis (Davies and Stone, 1995, Gopnik and Goldman, 1993) except to say that there are reasons for considering the utility of a simulation approach for the attribution of affective states. For instance, Adolphs (1999) points out that 'simulation' may be of particular use with concepts that can't be encoded lexically. This may include the qualitative component of an emotion including factors such as intensity as opposed to valence.

Work using physiological measures suggests that 'emotional contagion' may be reliant on a crude imitative process, which is consistent with a simulation account. This idea goes as far back as Charles Darwin (1872) who observed that people tend to subtly mimic the facial expressions of others in an emotional state. In fact, some authors define emotional contagion as the tendency to "automatically mimic and synchronise facial expressions, vocalisations, postures and movements with those of another person and consequently to converge emotionally' (Cacioppo *et al.*, 1988). More recently, Sonnby-Borgstrom, Jonsson and Svensson (2003) found that people high on empathy as measured by the Questionnaire Measure of Emotional Empathy (Mehrabian and Epstein, 1972 see chapter 2) showed 'stronger' facial motor mimicry reactions to targets as measured by electromyographic activity (EMG).

Several recent models of empathy from the field of affective neuroscience also have striking similarities with simulation. The 'shared manifold hypothesis' proposed by Gallese (2003) builds on recent evidence of the existence of 'mirror neurones' that fire in response to both observing and carrying out an action. Decety and Chaminade (2003) have also put forward a 'shared representations' model based on the premise that perception and action share the same neural and cognitive processes. They argue that perception of a behaviour in another person leads to activation of one's own representation of that behaviour and draw on evidence from neurophysiological and functional imaging studies of paradigms that manipulate imitation, observation and execution of action (see chapter 7 and Blakemore and Decety, 2001, Chaminade *et al.*, 2002, Decety and Chaminade, 2003, Decety *et al.*, 2002, Ruby and Decety, 2001). One study found behavioural interference when participants both executed and observed the same movements simultaneously (Kilner *et al.*, 2003). Decety and Chaminade (2003) suggest that this 'shared representations' model is consistent with what they call 'sympathy' – which we term 'affective empathy'.

Another similar model has been proposed, known as the Perception Action Model (Preston and deWaal, 2003). Although the authors give an account of affective empathy, they point out that the model focuses on process rather than response and is therefore broad enough to incorporate all forms of empathy. They summarise the model: 'an attended perception of the objects' state automatically activates the subjects' representation of the state, situation, and object, and that activation of these representations automatically primes or generates the associated autonomic and somatic responses, unless inhibited'[10]. A recent study found that participants imagining and recalling their own emotional experience and someone else's emotional experience activated the same brain areas (Preston, 2002).

[10] Damasio's (1994) somatic marker hypothesis is consistent with this model

Although the full philosophical debate on mental state attribution lies beyond this thesis, the similarities between simulation accounts and recent models of empathy from neuroscience are important. In order to explore empathy as it relates to emotional states, the possibility of a simulation account cannot be ignored and this and the corresponding models need to be empirically explored. The first step would be to develop a cognitive test of 'simulation' which does not rely solely on neuropsychological data.

3. Empathy: the neural correlates

Advances in methods for mapping functions to the brain have led researchers to propose candidate areas for localisation of the processes involved in social cognition. As far back as 1992 a social brain network was described comprising the orbito-frontal cortex, superior temporal gyrus and amygdala (Brothers and Ring, 1992). The medial frontal cortex has also been shown to be activated in functional magnetic resonance imaging (fMRI) paradigms using the standard false belief tasks adapted for the scanner (Fletcher *et al.*, 1995 and see chapter 7 for overview). These tests require the comprehension of first and second order false beliefs alone (cold cognition).

Evidence to support Brothers and Ring's model has been forthcoming. Stone *et al.*, (1998) tested people with bilateral damage to the orbito-frontal cortex on a task requiring the apprehension of a *faux pas*: a task tapping more complex emotion attribution. The volunteers were also given some first and second order ToM tasks. Although the participants could do the ToM tests ('cold cognition') – they had trouble with the *faux pas* despite being able to appreciate that the protagonist in the story would have been upset or hurt. The authors concluded that this brain area may be involved in the integration of mental state and affective information.

Farrow *et al.,* (2001) also found the orbito-frontal gyrus to be activated, along with the superior frontal gyrus and precuneus, in an fMRI study requiring participants to judge the emotional states of others', along with the forgivability of their crimes. Eslinger (1998) reviewing the functional data on frontal lobe lesions also found that the orbito-frontal systems may be implicated in affective empathy. Other studies also implicate this brain region with problems in social cognition (see Adolphs, 2001b).

There is also evidence to suggest that the amygdala is involved in mental state attribution (Adolphs *et al.*, 1999, Adolphs *et al.*, 2001c, Canli *et al.*, 2002) People with bilateral amygdala damage have trouble attributing complex mental states from pictures of the eyes alone (Canli *et al.*, 2002). Furthermore, our group has also found that people with amygdala damage have problems with the *faux pas* test (see above) and tests that tap the idiosyncratic emotional reactions stemming from complex beliefs (Shaw *et al.*, 200[illegible] in press.). The amygdala is also thought to be specialised for perceiving emotional expressions with negative valence[11] (Dolan *et al.*, 2001), although it seems to be involved in more general basic emotion perception as well. It is not clear, however, whether the amygdala is also recruited in the perception of non-facial social information such as body language (Adolphs and Tranel, 2003).

This raises the issue of whether neural circuits are specialised for perceiving certain emotional states (see chapter 7 for overview). As already mentioned, the amygdala is implicated in the recognition of basic emotions but may be especially implicated in fear and perhaps other emotions of negative valence and potential dangers in the

---

[11] There is some work suggesting it may not be valence specific and that past findings are due to inadequate controls for intensity between positive and negative emotions. For instance see Hamman, S., Ely, T., Hoffman, J. and Kilts, C. (2002) Activation of human amygdala in positive and negative emotion *Psychological Science*, **13**, 135-141.

environment [see above and Le Doux (1998)]. The processes involved in the perception of facial expressions of disgust, however, are thought to be localised in the insula (Philips, 1998), although again this area is also implicated in emotion perception *per se*. The idea that brain regions are functionally specialised for specific emotions (see chapter 7) is of importance when attempting to localise empathy (see Blair and Perschardt: response to Preston *et al.*, 2003). If the neural correlates of different emotions are fractionated along with different classes of mental state then localising empathy may involve many different brain regions.

More recent work has focused on the localisation of the simulation type models (see above and chapter 7). Decety and co-workers in a series of studies using observation and imitation paradigms (see above) have shown the premotor areas, right inferior parietal areas, precuneus, and somatosensory cortices to be involved in 'shared representations'. The right inferior parietal cortex is thought to be of particular importance in distinguishing self from other actions (Blakemore, 2003b, Blakemore *et al.*, 2003). The discovery of 'mirror neurones' in the premotor area F5 of the macque monkey also supports the idea that the premotor cortex is involved in simulation in humans (Gallese and Goldman, 1998). But some of this evidence is indirect, being based on the assumption that action imitation is a valid model for empathy, and so needs to be viewed cautiously. However, this assumption does seem consistent with recent findings that similar brain areas are activated (as those in action recognition) when viewing other peoples' mental states. Decety and Chaminade (2003) found similar activation when participants viewed film clips of people discussing emotional events, as did Carr *et al.*, (2003) when participants both imitated and observed others' emotional expressions (see chapter 7 for an overview). More evidence is needed of activation of the 'shared representation' network in mental state attribution paradigms, if this is to be of use as a model of empathy.

Adolphs *et al.*, (2000) used 3D lesion mapping on a sample of 108 people with focal brain lesions and found emotion perception to recruit the somatosensory-related cortices[12]. The authors point out that this is consistent with the idea that 'we recognize another individual's emotional state by internally generating somatosensory representations that simulate how the other individual would feel'. Other studies have also implicated somatosensory related areas in the imitation of emotional expressions. Carr *et al.*, (2003) found insula activation when people both observe or actively imitate emotional expressions, and comments' that the activation over both conditions was remarkably similar.

Another area that may be of importance for empathy is the anterior cingulate. Vogeley *et al.*, (2001) had participants adopt both the first and third person perspective in a traditional ToM task, and compared areas of brain activation. This area was implicated in both the attribution of mental states to the self and to other people (see chapter 7). This kind of study reminds us how brain areas and processes implicated in the representation of one's *own* mental states may be important in attributing states to others, in line with a 'shared representations' approach. Lane *et al.*, (1997) found the rostral anterior cingulate to be activated in a Positron Emission Tomography study, where participants were directed to attend to their subjective emotional responses. Critchley *et al.*, (2003) report work with people with focal damage to the anterior cingulate and healthy volunteers suggesting that the dorsal areas of this structure are

[12] This is also consistent with Damasio's work on the somatic marker hypothesis – see Damasio, A. (1994) *Descartes' error: Emotion, reason, and the human brain*, Avon Books, New York.

involved in the modulation of bodily arousal states. In a review paper, Mary Phillips and colleagues (2003) further implicate the anterior cingulate in the different facets of emotion perception.

It seems therefore that two broad accounts of neural circuits are emerging as candidates for mental state attribution and/or empathy. The first are the frontal areas primarily for epistemic states and the second, are 'body' (parietal, somatosensory and motor-related) area's for affective states. However, this distinction is not absolute and there is overlap in areas of activation between paradigms using 'cold' and 'hot' cognition. This is not a surprise: neither account is likely to be mutually exclusive, tasks are often complex and may invoke both strategies, and there may be individual differences and optimum conditions for the use of each approach. In addition, research on simulation circuitry in the brain is still fairly new, and the current data needs to be replicated and the new models thoroughly tested.

4. Empathy: conditions

Disruptions in the empathic process are part of the symptomatology of several psychological and developmental conditions. Clinical observation and diagnostic criteria lend support to the accuracy of these labels, but empathy has remained ill-defined with empirical data scarce and inconsistent. The following conditions have all been associated with empathy:

Psychopathy: The Psychopathy Checklist (Hare *et al.*, 1992) cites a lack of empathy as a key component of one of its three factors. Blair (1996) suggests this may be due to a deficit in the "mechanism which normally regulates the emotional responsiveness to distress signals ".

Williams Syndrome: Tager-Flusberg and Sullivan (2000) suggest that children with Williams Syndrome are impaired on false belief tasks but not on tasks requiring perceptual understanding of affective states. The authors argue for a dissociation on the basis of the type of mental state (see chapter 6).

Alexithymia: People with alexithymia tend to focus on external, objective facts rather than on their own feelings or those of others (Guttman and Laporte, 2002). It has been recognised as a clinical phenomenom in conditions such as borderline personality disorder and anorexia nervosa and has been found to be inversely related to empathy (Guttman and Laporte, 2002) as measured by the Interpersonal Reactivity Scale (Davis, 1980 and see chapter 2).

Schizophrenia: It is thought that ToM processes may play a role in some delusional states (Frith and Corcoran, 1996). People with schizophrenia have also been shown to have differential neural activation to controls when performing ToM tasks (Russell *et al.*, 2000).

Obsessive Compulsive Disorder/Anorexia/Paranoid Disorder(s): Gillberg (1992) proposes that these conditions be grouped together as disorders of empathy, as they can all be characterised by limited capacity for social interaction, extreme obsessiveness and alexithymia.

Turner's Syndrome: People with Turner's Syndrome lack a complete X chromosome: recent reports also indicate a possible deficit in emotion processing and perhaps empathy. Lawrence, Campbell, Swettenham, Terstegge, Akers, Coleman and Skuse (2003) found a group of girls with Turners Syndrome to be impaired on the 'Eyes task'

(Baron-Cohen *et al.*, 2001). This study also showed a deficit in a task requiring the identification of fear from the upper part of the face.

The two conditions that are to be the focus of this thesis are Asperger's Syndrome, a mild form of autistic spectrum disorder; and Depersonalisation Disorder, a condition characterised by an altered perception of the self, including one's own emotions.

C. 4.1 Autism spectrum disorders (ASDs)

People with autistic spectrum disorders suffer from a myriad of symptoms. The diagnosis for ASDs includes a reference to 'a qualitative impairment in social interaction' (DSM-IV, 1994 and see chapter 3 for full details of diagnosis). One of the symptoms of ASDs is a particular difficulty with many of the skills required to employ a ToM. Asperger's Syndrome (AS) is a milder form of ASD, and a diagnosis is given when someone meets the criteria for autism yet has normal intelligence and language skills (ICD-10).

The ToM account of autism is well established and refers specifically to the problems people with ASDs have with the *representational* aspects of ToM. People with AS also display deficits with this component of ToM, but to a lesser degree, and they tend to fail only the more complex second order and other advanced ToM tasks. However, the role in these tasks of domain general factors such as IQ and 'executive function' processes in these tasks still needs to be fully clarified (see chapter 6), and may have a bearing on the performance of people with AS on this tasks.

An alternative area of research has focused on the ability of people with ASDs to process affective mental states. Hobson, Ouston and Lee (1988) tested autistic children on a task requiring the matching of emotional voices with emotional expressions and found them to perform less well than the control group. However, the findings from studies using simple emotion expression paradigms are inconsistent (see chapter 6 for full overview). Bormann-Kischkel *et al.*, (1995) compared people with high-functioning autism (HFA) against a matched control group on a task requiring emotion recognition from photographs and the autistic group were significantly impaired. However, Loveland *et al.*, (1997) found no differences between controls and those with autism on emotion identification tasks using both verbal and non-verbal information. Blair (2003c) argues that such inconsistencies disappear once extraneous variables such as IQ are properly controlled, and that in fact people with ASDs do not have any problems appreciating facial affect. He then goes onto suggest that this profile i.e. proficiency with facial affect whilst finding 'cold' ToM cognition difficult, is opposite to that observed in people with psychopathy, hence providing evidence of a dissociation of function (Blair, 2003c).

Children with ASDs have also been shown to have imitation deficits (Williams *et al.*, 2001) which may be a precursor to their more extensive ToM problems (see chapter 5 for overview). For instance, one study found that a group of very young children with autism (i.e. 20 months old) displayed less imitation than comparison groups (Charman *et al.*, 1997 - and see chapter 5 for overview). Stone *et al.*, (1997) also found that although ~~of~~ autistic children showed weaker imitation skills, they showed a similar pattern of performance across tasks, with imitation of body movements being more difficult for them than imitation with objects.

A few tests have also examined emotion attribution in ASDs using more complex designs. Baron-Cohen *et al.*, (1991) found that children with autism could understand emotions stemming from situations but not from beliefs. Autistic children were also quite good identifying a character's emotional state from a video and demonstrating a parallel emotional state (Yirmiya *et al.*, 1992). Although they were significantly worse than age and gender matched controls, they did surprisingly well. People with ASDs have also been found to have intact physiological responses to both non-person based and person-based emotional stimuli, as measured by galvanic skin response (Blair, 1999).

People with AS seem to do badly on more complex ToM tasks. Baron-Cohen (1997) found that despite passing traditional ToM tasks, people with AS and HFA had trouble with appreciating a difficult *faux pas*. Such tests tap the ability to integrate inferential ToM with affective information (see above). Furthermore, Baron-Cohen *et al.*, (2001) tested people with AS and HFA on their ability to attribute mental states simply from a picture of a person's eyes, and found that they performed significantly worse than controls. Channon *et al.*, (2001) examined adolescents with AS on a test requiring the appreciation of real-life problem-solving and solution generation using situations such as how to deal with noisy neighbours. The group were found to be impaired on several indices.

The ability of people with AS to empathise with others' emotional states both intellectually and emotionally is still under debate. This is in part due to the fact that those with HFA and AS are often combined to form one group making it difficult to ascertain the specific deficits of those with AS[13]. Shamay-Tsoory (2002) gave self-report empathy questionnaires, the *faux pas* task and emotion processing tasks to 2 people diagnosed with Asperger's Syndrome, and found that they scored lower than

[13] Some argue that the distinction between high-functioning autism and AS should be dropped.

controls. The author proposes that the deficit in AS lies specifically in integrating cognitive and affective material.

Baron-Cohen and Wheelwright (2004 in press) also found that those with HFA and AS scored significantly lower than controls on a self-report 'Empathy Quotient' which taps all facets of empathy (see chapter 2). However, items tapping cognitive and affective empathy were lumped together, making it difficult to tease out the groups' affective empathy skills. It also remains unclear how people with AS perform on paradigms tapping imitative ability. One study found a group of 4 people with AS to have intact function of 'mirror neurones' in an object manipulation task – but this may have been an artefact of sample size (Avikainen *et al.*, 1999). In addition, the sparing of language in this group may have important implications for mental state attribution, as links have been made between the brain areas responsible for these two functions[14] (see chapter 5).

So it seems that, in contrast to those with more severe forms of autism, people with AS are *relatively* spared the 'cold' aspects of ToM, although it not clear whether their performance is qualitatively the same. However, there are few data on the ability of people with Asperger's Syndrome to appreciate affective states that require 'hot' cognition. This issue is potentially interesting in view of the fact that different types of empathy may be dissociated both cognitively and neuroanatomically, and that such knowledge could play a vital role in interventions if necessary.

[14]There has been a recent suggestion that the traditional language region of the brain may be involved in imitation Heiser, M., Iacobobi, M., Maeda, F., Maruc, J. and Mazziotta, J. (2003) The essential role of Broca's area in imitation *European Journal of Neuroscience*, **17**, 1123-1128.

4.2 Depersonalisation disorder (DPD)

This condition is defined as an 'alteration in the perception or experience of the self so that one feels detached from and as if one is an outside observer of one's mental processes or body' (DSM IV, 1994). People suffering from depersonalisation disorder (Senior *et al.*, 2001) frequently complain of a subjective deficit in empathising. However, to our knowledge, despite these anecdotal reports, there are no studies examining empathic ability in DPD.

Theoretically, several possible explanations may account for these subjective reports. People with DPD are known to show problems in emotion processing. Emotional blunting is often reported in clinical settings and has been verified empirically. Sierra *et al.*, (2002) found that, as a group, people with DPD showed less physiological activation as measured by galvanic skin response (GSR). In addition, Phillips *et al.*, (2001) found people with DPD rated scenes to be less aversive than a control group, and that they failed to activate the same brain regions whilst viewing these. It is entirely plausible that this problem in processing emotional material also affects the attribution of emotional states in others. This is especially possible if one accepts a simulation/shared representations account of empathy that relies on the use of the self-knowledge to interpret other people's affective states. The empathy deficits reported by this group may therefore be an objective by-product of differential emotion processing.

On the other hand, an objective empathy deficit could exist for entirely different reasons. People with DPD often have a co-morbid diagnosis for depression and anxiety (Lambert *et al.*, 2001). It is therefore conceivable that either of these conditions (or their complex interaction) could affect empathic ability, perhaps via low level emotional reactivity (anxious or depressed mood) or via motivational factors as in the case of depression. The relationship between DPD, depression and anxiety is particularly hard to disentangle and it is also possible that sufferers of DPD are not a homogenous group in this regard.

A final possibility stems from cognitive behavioural accounts of DPD. Cognitive models of anxiety, depression and panic include processes such as negative attribution bias and a tendency for negative (sometimes catastrophic) evaluation of the self and/or situation. Recent DPD models have drawn on this literature and suggested that similar processes may be at work in DPD (Hunter *et al.*, 2003). Such negative evaluations could lead the sufferer to incorrectly evaluate their own ability to empathise and interact with others. This would suggest that the reported empathy deficits are subjective rather than objective.

In order to test these claims, it is necessary to administer both self-report and objective tests of empathy and perform statistical comparisons between the two. This will shed light on the anecdotal reports of this distressing lack of fellow-feeling reported by some sufferers of DPD and, in turn, contribute to treatment interventions.

5. Aims of the thesis

Standardised ToM tests, self-report measures and novel experimental tasks will be used to ascertain whether people with AS and DPD have objective empathy deficits, and to further clarify the nature of any problems in this domain.

The expectation is that people with AS will display evidence of difficulty with the more advanced tests of ToM which focus on intentional and epistemic states. However, it is unclear whether such marked problems will emerge in tests that tap the ability to appreciate other peoples' affective states (with and without emotional resonance). We will therefore employ novel tasks to explore whether there is a dissociation in this clinical group between empathy measures as a function of the type of mental state.

Both novel tasks and self-report measures will also be used to explore whether people with DPD have objective or subjective disturbances in their ability to understand and resonate with the emotional states of others. From anecdotal reports alone it is not clear whether the reported lack of empathy results from an objective empathy disturbance, a disruption of attributions about the self, or from the often co-morbid symptoms of depression and anxiety.

A novel fMRI paradigm will also be used to explore the neural correlates of simulation (likely to be a key process in empathy – see above) in healthy volunteers. There is very little direct empirical research into the neural correlates of this process. It is hypothesised that a simulation strategy will activate the pre-motor and other body-related areas such as the parietal cortex, as opposed to the brain regions that are activated when attributing non-affective mental states or during 'cold' cognition.

Measuring Empathy: Self-Report Measures

1. Overview of self-report measures

In order to measure empathy, it is useful to have a reliable and valid self-report scale. Although objective measures are important, people often have some idea of their own empathic capacity and it is important to utilise this insight when attempting to assess empathy levels. Several scales have been developed over the years, each reflecting the different conceptualisations of empathy. However, they tend to be either based on limited definitions or to have other important weaknesses.

Contents at a glance

1.1 Questionnaire Measure of Emotional Empathy (Mehrabian and Epstein, 1972)

This Questionnaire Measure of Emotional Empathy (QMEE) was designed to tap emotional empathy, which the authors define as a 'vicarious emotional response to the perceived emotional responses of others'. It includes items such as:

- I get very angry when I see someone being ill-treated.
- I am very upset when I see an animal in pain.
- I cannot continue to feel OK if people around me are depressed.
- It makes me sad to see a lonely stranger in a group.

In an effort to control for participants responding in a socially desirable way, items were taken from a larger set on the basis of non-significant correlations with the Crowne and Marlowe (1960) Social Desirability Scale (SDS). This process is necessary to address some of the criticisms of self-report measures, which in this case, would show as people wanting to appear more 'empathic', empathy being a socially desirable trait. The final version of the QMEE has a correlation of $r = 0.06$ with the SDS.

Factor analysis of data gathered from the QMEE revealed several underlying clusters, such as susceptibility to emotional contagion, appreciation of the feelings of unfamiliar and distant others, emotional responsiveness, and a tendency to be moved by others' positive/negative experiences. The scale is deemed to have good internal reliability: split-half reliability is .84 and individual items all correlated at $p = .01$ level with total score on scale.

Despite the scales' reliability, it is not clear that it actually measures empathy. Construct validity was measured on the basis of prediction of aggressive behaviour – which may or may not be directly related to empathy. In fact, the authors suggest that it may measure general emotional arousability as opposed to reactions specific to others' emotion (Mehrabian *et al.*, 1988b). This eventually led them to construct a new scale, the Balanced Emotional Empathy Scale (Mehrabian *et al.*, 1988a) - which is limited more to reactions specific to others' emotions. Unfortunately, it is again unclear whether the items tap only emotional empathy, for example, 'I cannot easily empathise with the hopes and aspirations of strangers/I easily get carried away by the lyrics of a love song'.

1.2 The empathy scale (Hogan, 1969)

This measure was also developed in the 1960's and purports to measure cognitive empathy, including role-taking and perspective-taking. The authors define cognitive empathy as 'the apprehension of another's mental state - without actually experiencing that state'. It includes items which deemed indicative of an empathic response if answered true:

- I am not easily angered. (true)
- I must admit I often try to get my own way regardless of what others may want. (false)
- It is the duty of a citizen to support his country. (false)
- As a rule I have little difficulty putting myself in other peoples' shoes. (true)

Unfortunately, no items were included to control for socially desirable responding, however, test reliability was good at .84. In terms of construct validity, the questionnaire correlates significantly with scales measuring communication competence, moral maturity, social insight and effective social functioning. The initial factor analysis revealed three underlying dimensions, these being even-tempered disposition, sociable interpersonal style and humanistic socio-political attitudes.

However, a later factor analysis suggested it may actually measure social self-confidence, even temperedness, sensitivity and nonconformity (Johnson *et al.*, 1983). In criticism, the emergence of a factor tapping attitudes and the inclusion of items such as 'it is the duty of a citizen to support his country' (a negative response is deemed empathic) suggest the scale relies upon assumptions about the consequences of being empathic which may be influenced by cultural norms as opposed to measuring empathy *per se*. It has also been suggested that the scale simply measures social skills rather than pure empathy (Davis, 1994). While it is true to say that social skills rely on cognitive empathy to a degree, they may also be reliant on other skills such as affective empathy, the ability to integrate emotional and cognitive states, emotion perception and possibly an element of rule-based reasoning.

1.3 The interpersonal reactivity index (Davis, 1980)

The Interpersonal Reactivity Index (IRI) is a multi-dimensional measure of empathy which includes several constructs not considered in previous self-report scales. The author criticises preceeding scales (see above) for summing both cognitive and emotional aspects of empathy into one single score. The IRI includes subscales that measure: perspective-taking which fits with traditional definitions of cognitive empathy; empathic concern (which specifically addresses the capacity of the respondent for warm, concerned, compassionate feelings for others i.e. a facet of affective empathy); fantasy items which measure a tendency to identify with fictional characters; and personal distress (which is designed to tap the occurrence of self-orientated responses to others' negative experiences - see chapter 1 for discussion).

Factor analysis confirmed the presence of each of these of underlying dimensions. Each subscale consisted of 7 items each measured on a 5 point Likert scale ranging from 'does not describe me at all' to 'describes me very well'. The cognitive items include such questions as:

- Before criticising someone, I try to imagine how I would feel if I were in their place.
- I sometimes try to understand my friends better by imagining how things look from their perspective.

Examples of empathic concern items are:

- I am often quite touched by things that I see happen.
- When I see someone taken advantage of, I feel kind of protective towards them.

Fantasy items on the other hand focus on imaginative capacity:

- I really get involved with the feelings of the characters in a novel.
- I daydream and fantasise, with some regularity, about things that might happen to me.

Lastly, the personal distress subscale includes items such as:

- In emergency situations, I feel apprehensive and ill-at-ease.
- When I see someone who badly needs help in an emergency, I go to pieces.

The questionnaire has high test-retest reliability between $r = .61$ and $r = .81$. and internal reliability ranges from .75 to .78 for each of the subscales. But importantly the

IRI does not control for social desirability. Furthermore, although Davis (1980) claims to be measuring empathy as a global concept, others have pointed out that the fantasy items do not tap pure empathy (Baron-Cohen and Wheelwright, 2004 in press) and despite being crucial to its measurement, personal distress is not classed as empathy. While it may be useful to use the subscales in isolation (see Davis, 1994) – as each subscale only includes 7 items – the range of scores would be greatly reduced.

1.4 The empathy quotient (Baron-Cohen and Wheelwright, 2004 in press).

The Empathy Quotient (see appendix 1) is the most recent addition to self-report measures of empathy. Unlike previous scales it was explicitly designed to have a clinical application and be sensitive to a lack of empathy as a feature of psychopathology. There are 60 items including 20 filler items; responses are given on a 4 point scale ranging from 'strongly agree' to 'strongly disagree'. Approximately half the items are reversed. Participants receive 0 for a 'non-empathic' response, whatever the magnitude, with 1 or 2 for an 'empathic response' depending on the strength of the reply and so the total score is out of 80. Six experimental psychologists working in the field independently rated each item as to its association to a given definition of empathy.

The EQ was validated on 197 healthy control volunteers and 90 people with Asperger's Syndrome (AS) or high-functioning autism (HFA) and age and gender matched controls (a gender ratio of 2.6:1 m:f was found). It was shown reliably to distinguish between the clinical and control groups. Consistent with previous literature, the authors found gender differences in the control group, with women scoring significantly higher. In addition, the EQ was found to be correlated with a

Friendship Questionnaire designed to measure and intimacy in relationships (Baron-Cohen and Wheelwright, 2004 ~~in press~~). The EQ was also shown to have high test re-test reliability (r= 0.97) over a period of 12 months. Furthermore, Baron-Cohen *et al.*, (2003b) recently replicated the female superiority on the EQ and showed once again that it distinguished between those with AS/HFA and controls.

As the EQ covers both cognitive and affective empathy using a wide range of items and has been validated on clinical groups, it seemed the most appropriate scale for use in this study. As it is fairly new, information regarding the scales construct validity is limited. It was therefore necessary to further examine the reliability and validity of the EQ prior to its use.

2. The empathy quotient: reliability and validity[15]

The EQ was therefore given to several different groups in order to check its validity and utility across samples. Test-retest reliability was re-examined, as was the association between the EQ and various other 'empathy' measures in order to examine the scale's validity. Specifically, the relationship between the EQ and the Eyes task (Baron-Cohen *et al.*, 2001) was explored as the latter is an objective measure of cognitive empathy and hence an indicator of construct validity. The association between the EQ and the Interpersonal Reactivity Index (IRI Davis, 1980) was also considered, as a check on concurrent validity. Furthermore, to our knowledge the inter-relation between the EQ items and social desirability has not yet been examined. This knowledge is necessary in order to address a general problem with self-report measures: that people may respond according to how they would like

[15] This chapter is in part reproduced from Lawrence, E., Shaw, P., Baker, D., Baron-Cohen, S. and David, A. S. (2004 in press) Measuring empathy - reliability and validity of the Empathy Quotient, *Psychological Medicine*, **34**.

to appear, that is to say, highly 'empathic'. Lastly, an exploratory factor analysis was performed in order to consider the factor structure of the scale and to shed light on the various components of empathy.

2.1 Construct validity

Participants

There were 53 volunteers consisting of 28 women and 25 men with a mean age of 32.5 years (± 10.9). Approximately, half of this group were recruited from mental health professionals at the Institute of Psychiatry, London. The remainder were recruited from non-academic/clinical staff and through advertisements in the local area.

Procedure

All measures were administered in a quiet room as part of a wider testing session. Participants were given the EQ (Baron-Cohen and Wheelwright, 2004 ~~in press~~) self-report measure of empathy. Missing values on the EQ, resulting from a double endorsement or no endorsement, were substituted with the group mean rounded to the nearest whole number.

Participants were also given the Social Desirability Scale (SDS - Crowne and Marlowe, 1960). On this measure, 1 point is allocated for each item endorsed, resulting in scores ranging from 1 to 33 with a high score indicating that the respondent is prone to give answers in accordance with social desirability, e.g. 'I sometimes feel resentful when I don't get my own way'.

The Eyes Test (Baron-Cohen *et al.*, 2001) was also administered. This test measures the ability to decipher a mental state from a picture of the eyes alone, and according to the authors, is an advanced measure of 'mind-reading' (see appendix 2). In terms of the various components of empathy, this task would be expected to tap cognitive empathy and it has the advantage of not relying on self-report. This test has been shown to distinguish reliably between people with AS/HFA and healthy controls. One point is allocated for each correct answer with the final score being out of 36.

Lastly, participants completed the National Adult Reading Test (NART - Nelson, 1982). Participants read 50 irregular sounding words (i.e. ache) which yields an estimate of verbal intelligence.

Results

Mean EQ scores for all items for both men and women can be found in Table 2.1 (see below). These are similar to those found in the original study (Baron-Cohen and Wheelwright, 2004 ~~in press~~) i.e. males 41.8 (±11.2) and females 47.2 (±10.2). Furthermore, the difference between men and women was again significant ($t = -3.5$, df 51, $p = .001$). The data were normally distributed [slightly negative skew (-.190) and kurtosis of less than 1 (-.717)].

Table 2.1 Mean and SD EQ scores stratified by gender

		total score on the EQ				
		n	Mean	SD	Min	Max
gender	male	25	41.28	10.1	22	58
	female	28	50.57	9.2	30	66
Group Total		53	46.19	10.6	22	66

Each item on the EQ was entered into a Pearson's Product Moment Correlation analysis along with the total score on the SDS. A positive correlation above .3 was taken as an indicator of socially desirable responding. Items 11, 18, 27, 34, and 37 of the EQ, all correlated significantly with total SDS score, but item 27 correlated below .3, and item 37 had a negative rather than positive relationship. Items 11, 18 and 34 were therefore dropped from subsequent analyses.

The mean score on the Eyes Test was 27.6 (± 4, n = 48), again very similar to the normative data. In the original sample, participants from the general population had a mean of 26.2, and participants drawn from a sample comprising students had a mean of 28. These data were then correlated with total EQ score and a modest positive relationship was found (n = 48 r = .294, p = .033).

The estimated IQ score from the NART for this group was 120.48 (± 4.7), which is above normal range. As both the Eyes Test and EQ have verbal components, a correlational analysis was run to examine the association between verbal IQ, as estimated from the NART, and each of these variables. There was a near significant relationship between performance on the Eyes Test and verbal IQ (r = .385, p = 0.07) but not between the total EQ score and verbal IQ

As a result of this relationship a multiple regression analysis was performed to include both total EQ score, verbal IQ and other demographic factors (gender, age, education and whether the participant was a clinician/academic or not). The only significant predictor of the Eyes test was verbal IQ score (multiple $R = .369$), which accounted for 11.7% of the variance. However, both gender ($r = .266$, $t = 1.83$, $p = .074$) and EQ score ($r = .255$, $t = 1.75$, $p = .087$) approached significance.

Discussion

Scores on the EQ did not deviate substantially from the Baron-Cohen data, indicating the scale's reliability across samples. There were also significant differences between men and women, as in the original data. Women scored slightly (but not significantly) higher than in the original sample, which may be because a higher proportion of them were drawn from mental health workers. This may prevent firm conclusions regarding gender differences.

Five items correlated significantly with total score on the SDS and 3 of these were weak to moderate positive correlations, hence were left out of the analysis. The negative correlation with item 37 is somewhat mysterious and may be due to chance factors. That the remaining 35 items showed no association with social desirability supports the scale's construct validity.

The mean score on the Eyes Test matched those found by the original authors confirming a successful replication. As expected, this task was significantly associated with total score on the EQ. This is of particular interest, as the Eyes Test provides an objective measure of empathy and hence adds further weight to the validity of the EQ. However, the fact that this association did not reach significance in the regression analysis needs to be considered. Verbal IQ, as estimated by the NART, was the sole predictor of performance on the Eyes Test. This raises the possibility that the EQ and Eyes task tap different constructs. Nevertheless the weak but non significant partial correlation with total EQ score and the independence between verbal IQ and EQ suggest that perhaps performance on the Eyes task is reliant on *both* verbal IQ and empathy.

An alternative explanation for the role of verbal IQ is that it is an artefact of sample selection. A selection bias could have led to a particularly high IQ range which may have been a confound: clinicians/academics in the field of psychological medicine have high IQ's, and unrelated to this, they also do well on the Eyes Test. However, a dichotomous occupation variable measuring this occupational variation was not a predictor in the multiple regression. These results suggest that the relationship between the Eyes task and the different factors underlying the EQ, as opposed to the total score, should be investigated.

2.2 Concurrent validity and test-retest reliability

Participants

44 people were re-contacted between 10 and 12 months after Study 1. 25 people replied consisting of 9 men and 16 women with a mean age of 32 years (± 9.5). There were no significant age differences between this group and those reported in Study 1 ($t = 1.29$, df 51, $p > .05$) nor on gender distribution ($\chi^2 = 1.99$, df = 1, p = > .05). A further 4 people also completed one of the measures.

Procedure

Participants were sent both the Empathy Quotient and the Interpersonal Reactivity Index (Davis, 1980). The EQ was sent again in order replicate the test-retest reliability observed in the original study. The IRI was sent as a measure of concurrent validity. The range of scores for each subscale on this questionnaire is 0 – 35, with 35 representing a high 'empathy' score for perspective-taking, empathic concern and fantasy items subscales. A high score on the personal distress scale, however, is indicative of a tendency to display self-orientated emotional reactivity.

Twenty-five people filled out and returned the EQ. Twenty-four of this group also returned the IRI along with a further 4 people who filled out the EQ at time 2 only. This group therefore comprised 11 males and 17 females (mean age 32.1 years, ± 9.7). There were no age differences between this group and the participants in Study 1 (t = 1.03, df 124, $p > .05$), nor was there any difference in gender distribution ($\chi^2 = 1.22$, df1, $p > .05$).

Results

The test-retest correlation coefficient between EQs administered at time 1 and at time 2 was ($n = 25$, $r = .835$, $p < .001$).

A correlation matrix was generated to examine the relationship between the IRI and the total EQ score from questionnaires sent out at time 2. The 3 items that correlated with the SDS in Study 1 were again dropped from the analysis. Moderate correlations were found between the EQ and both the 'empathic concern' subscale ($n = 28$, $r = .423$, $p = .025$) and the 'perspective-taking' subscale ($n = 28$, $r = .485$, $p = .009$). The correlation coefficient was close to zero for the 'fantasy' items ($n = 28$, $r = -.027$, $p > .05$) and not significant for the 'personal distress' items ($n = 28$, $r = -.158$, $p > .05$).

Discussion

The EQ shows high test-retest consistency over a period of 10 to 12 months, a finding which adds yet more support for the reliability of this measure.

Furthermore, the relationship between the EQ and subscales on a previously validated self-report measure of empathy supports the scale's validity. The fact that the correlations are only moderate is as expected, as total EQ score is an index of global empathy. There was a significant relationship with both empathic concern and perspective-taking, indicative of concurrent validity. The lack of a significant relationship with 'personal distress' is also encouraging, as this is not intended to be a measure of empathy but of 'emotional reactivity'. The fact that a weak negative association was found supports the idea that 'personal distress' and empathy are inversely related. The absence of an association between the EQ and the IRI fantasy items is consistent with previous criticisms that this concept should not be considered a component of empathy (Baron-Cohen and Wheelwright, 2004 ~~in press~~).

2.3 Factor structure

Participants

An additional 57 volunteers [22 men (38.6%) and 35 women (61.4%)] completed the EQ. These participants were recruited by the author and Dr Philip Shaw (Institute of Psychiatry) during the course of other projects. These data were combined with those from Study 1 to create a control group of 110 psychologically healthy participants.

Furthermore, 54 people who contacted the Depersonalisation Research Unit at the Institute of Psychiatry, London, reporting symptoms of depersonalisation disorder (DPD), were also given the EQ. DPD is defined as an 'alteration in the perception or experience of the self so that one feels detached from and as if one is an outside observer of ones mental processes or body' (DSM IV, 1994). DPD can also include derealisation 'detachment from one's surroundings' and de-affectualisation, that is to say, 'emotional numbing'. DPD is also often accompanied by subjective complaints of a lack of empathy; it is unclear whether this is an objective deficit or not (see chapter 1 and 3). However, there is no reason to suspect a problem in employing cognitive empathy, so it is unlikely that this group have a global empathy deficit.

Some of these participants are a subgroup of a cohort reported elsewhere ~~(Baron-Cohen and Wheelwright, 2004 in press~~ Baker et al, 2003). A further 8 people diagnosed with DPD were also recruited from the same unit. As a whole, this group comprised 32 men (51.6%) and 30 women (48.4%), with a mean age of 34.6 (± 10.8). A chi-squared analysis revealed that the gender distribution was not significantly different between the control group and the DPD group ($\chi^2 = 1.26$, df1, $p > .05$). Neither was mean age significantly different between these two groups ($t = -.593$, df 113, $p > .05$). After some preliminary analysis, all the groups were combined ($n = 172$) to create a group consisting of 79 men and 93 women with a mean age of 34.1 years (± 10.4).

Measures

The EQ was sent both to the additional 57 people recruited for the control group and those that formed the DPD group. The DPD group were also sent some further mental health screening measures, such as the Dissociative Experiences Scale version II (DES - Bernstein and Putnam, 1986, Carlson and Putnam, 1993) and the Beck

Anxiety and Depression Inventories – BAI and BDI (Beck *et al.*, 1988a, Beck *et al.*, 1988b).

The DES is considered the 'gold standard' in measurement of DPD symptomatology. It is a 28 item self-report questionnaire with a cut-off score of 30 for severe dissociative disorders (Carlson & Putnam, 1993). Factor analysis has revealed three main components of the scale: 'depersonalisation/derealisation (DPD/DR)', 'amnesia' for dissociative experiences, and 'absorption' and imaginative involvement (Carlson *et al.*, 1991). Eight items make up the DES-Taxon which is sensitive to the detection of DPD with a cut-off score of 13 (Simeon *et al.*, 1998).

The BAI and BDI were also given to participants, due to the co-morbidity observed between depersonalisation disorder, depression and anxiety (Simeon *et al.*, 1998). A score below 11 on either scale is considered within the 'normal' range, whereas a score above 30 is classed as 'severe'.

Preliminary Analysis

The mean EQ scores for the DPD group (including all the items) can be found in Table 2.2. No significant differences were found in EQ score between the psychologically healthy group tested in Study 1 and those with DPD: for men ($t = .783$, df 55, $p > .05$) or women ($t = 1.485$, df 55, $p > .05$). The difference between men and women on total EQ scores again reached significance ($t = -.2.686$, df 59, $p = 0.009$).

Table 2.2 Mean and SD EQ scores for the ~~EQ~~ OPO group stratified by gender

		total score on the EQ				
		N	Mean	SD	Min	Max
gender	male	32	38.94	12.4	15	66
	female	30	46.76	10.1	23	65
Group Total		62	42.66	11.9	15	66

53 people completed the DES, BAI and BDI. The mean score on the BAI was 21.6 (± 12), and BDI was 20.3 (± 10.5). The mean score on the DES was 23.2 (± 14.2), the DPD/DR subscale 36.6 (± 24), amnesia 6.2 (± 2.5) and absorption 27.6 (± 17). The mean score on the DES-taxon was 23.3 (15.4).

The BAI, BDI and EQ were all entered into a correlational analysis and the coefficients were found to be close to zero. A correlation matrix for the EQ and the mean DES score along with the scores for the 3 subscales was also generated. No correlations reached significance, although there was a trend for a negative relationship between total EQ score and the DES amnesia subscale ($r = -.256$, $p = .067$ for a two-tailed test).

This analysis shows that people reporting DPD who suffer mild to moderate levels of anxiety and depression do not suffer a global empathy deficit. In fact, they show the same pattern of results on the EQ as the psychologically healthy population detailed above, including gender differences. However, it is entirely possible that people with DPD may display differences on the specific components of empathy, that is to say, they may have problems experiencing affective empathy due to the emotional blunting they often report. It was hypothesised that there would not be any difference between the EQ factor structures of the DPD group and healthy controls – despite any possible differences in their scores (the latter will be examined in chapter 3). Two separate factor analyses were conducted to test this prediction.

Procedure

An exploratory factor analysis, using a principal components analysis to construct the initial model, was performed on the EQ. Although the data are ordinal, many authors feel that this procedure is still useful as long as meaningful factors are extracted (Hutcheson, 1999). Gorsuch (1974) points to a problem with using non-continuous data in factor analysis: it can result in spurious factors where items load according to 'difficulty', and that although this must be considered when interpreting the factors, it is much less likely with trichotomies. Kim and Mueller (1978) argue that the main artefact of using ordinal data in exploratory factor analysis is that the factors may be harder to interpret. These issues were borne in mind when interpreting the data.

Nine cases had missing values ranging from 1 to 4 and were dealt with as described above. However, one additional participant had a whole page missing and so these values were left as missing.

Results

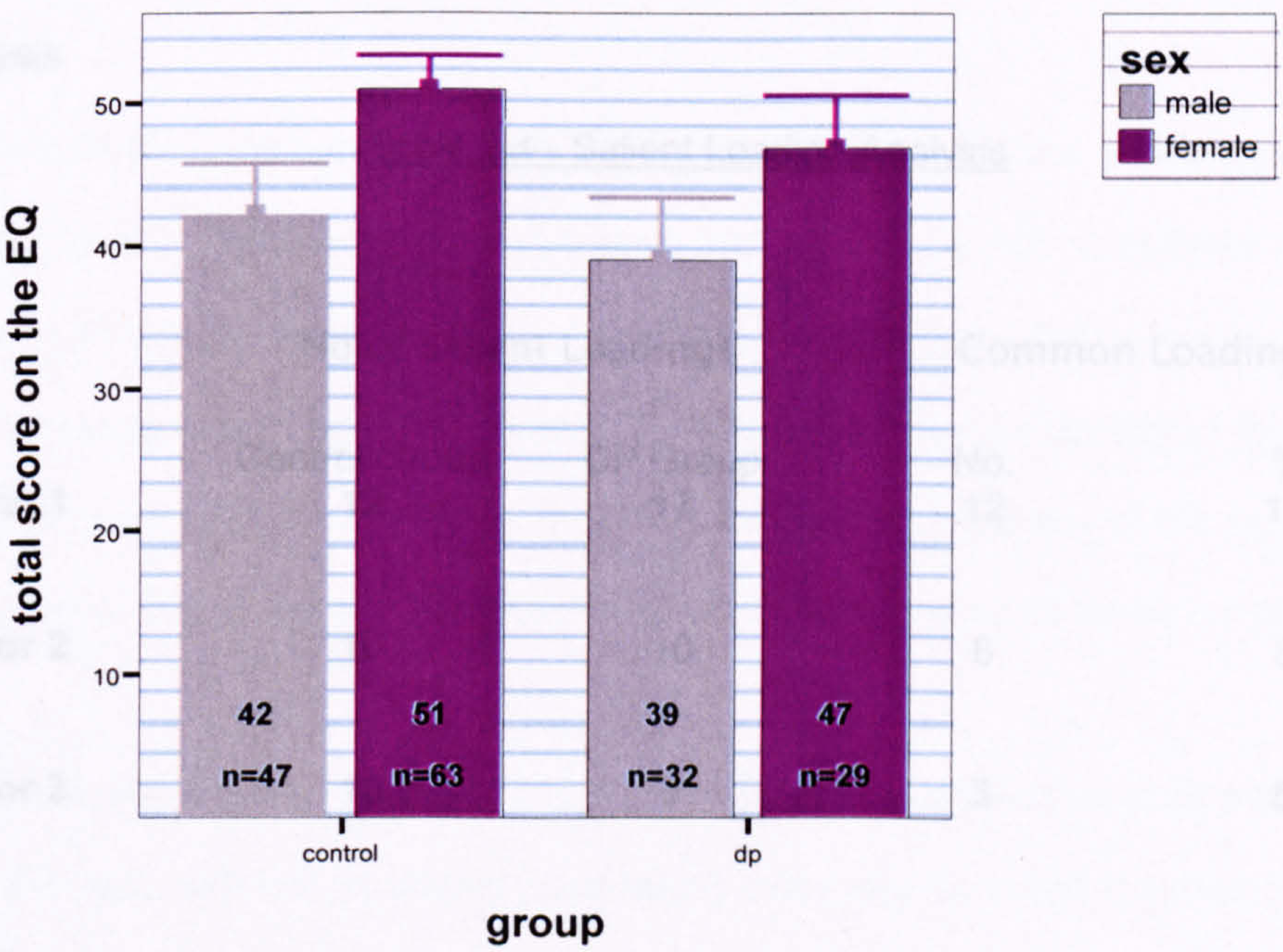
The mean and standard deviation EQ scores for the entire group can be seen in Table 2.3.

Table 2.3 Mean and SD EQ scores for entire sample stratified by gender

		total score on the EQ				
		N	Mean	SD	Min	Max
gender	male	79	40.9	11.9	15	66
	female	93	49.6	9.6	23	69
Group Total		172	45.58	11.6	15	69

These are remarkably similar to the normative data for both men (mean 41.8 ± 11.2) and women (47.2 ± 10.2). There were also significant gender differences between men and women ($t = -5.34$, df 147.38, $p = 0.001$) and this held both the control and DPD (see figure 2.1) groups mirroring the comparison made above between participants from study 1 and those with DPD.

Figure 2.1: Mean EQ scores for both men and women in the control and DPD Groups



Group Comparison – a separate analysis was conducted for each group (DPD vs. healthy volunteers) in order to examine the similarity of the factor structure as hypothesised. A principal components analysis followed by an exploratory factor analysis was then performed (as there was no *a priori* hypothesis about the underlying structure) with a varimax rotation to maximise the independence of the components. For each analysis a scree plot was used (Cattell, 1966) to decide upon the number of factors, as opposed to eigenvalues which may give rise to many uninterpretable factors. Values less than .3 were suppressed, since they accounted for less than 10% of the variance for that factor.

In order to compare the factor structures of the two groups, a salient loading profile (Abdel-Khalek *et al.*, 2002) was performed using .35 as a cut-off point to define a salient loading (see table 2.4). These figures were considered along with tentative labels for each of the factors for each group (Tabachnick and Fiddell, 1989). On the basis of this information the decision was made to combine all the data into one analysis.

Table 2.4 - Salient Loading Analysis

	No of Salient Loadings		Common Loadings	
	Control Group	DP Group	No.	%*
Factor 1	12	17	12	100
Factor 2	10	10	8	80
Factor 3	10	9	5	50

* the percentages were calculated in proportion to the control salient loadings

Data Screening – A Pearsons correlation matrix was generated, and all EQ items that failed to correlate with any other item at .2 (Hutcheson, 1999) or that had low communalities (i.e. a small amount of the item's variability was accounted for by the factor), in the final model were removed. This resulted in the removal of 8 items, namely 15, 18, 28, 37, 38, 39, 49, 60 (see appendix 1). Conversely, none of the items correlated too strongly, indicative of an absence of multicollinearity.

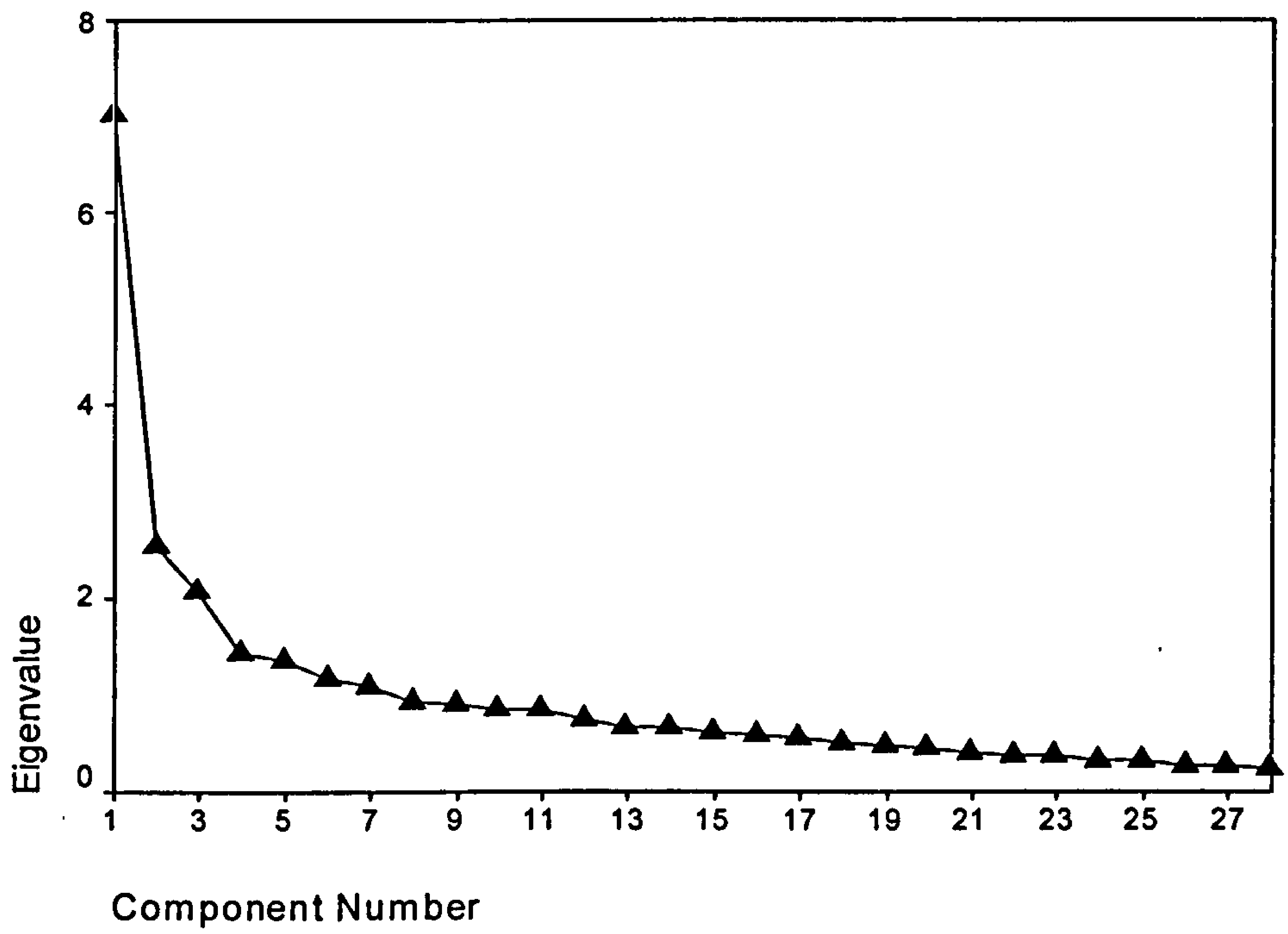
All the items were also correlated with the SDS. Five EQ items showed a significant relationship with the total score on this scale, namely 11, 18, 34, 37, 46. Again, item 37 showed a negative relationship; however, it also had a low loading, as did item 18 (see above), and so this stage of data screening only resulted in the removal a further 3 items. Eleven items were therefore left out of the analysis.

Sample Size – there were 29 items and 172 cases, conforming to the 5 cases per item rule for an adequate sample size for this kind of analysis.

Final Analysis - the scores for the remaining 29 items for the entire sample were subject to an initial principal components analysis (PCA) with a varimax rotation in order to assess the number of underlying factors in the data. The communalities lie in the mid range except for items 10 and 57. After examining the content of these, it was decided to keep 57 but to remove 10 which did not load onto any of the factors in the final model. The analysis was re-run without this item leaving 28 items in total.

The scree plot showed that only 3 or 4 plots (factors) appeared stacked and separate from the rest, with the remaining plots falling away and bunched together (see figure 2.2 below). The eigenvalues for the 4 factors were 6.95, 2.56, 2.082, and 1.45. The decision was made to keep 3 factors, as it was apparent from both the scree plot and eigenvalues that the first 3 are by far the strongest and they all had eigenvalues over 2. These factors account for 41.4% of the total variance.

Figure 2.2 Scree Plot for Entire Dataset



The item loading for these 3 factors in the rotated solution are shown in Table 2.5. Double loadings were allocated on the basis of content, with agreement reached between the two raters. The Keiser-Meyer-Olkin measure of sampling adequacy was 0.846 and the Bartlett test of sphericity was highly significant, suggesting the data were suitable for PCA. Factor 1 was labelled 'cognitive empathy', factor 2 'emotional reactivity' and factor 3 'social skills' - these labels will be expanded upon in the general discussion.

Table 2.5: Final Item Loadings from the Principal Components Analysis

	1	2	3
EQ55	.763		
EQ52	.726		
EQ25	.723		
EQ54	.696		
EQ44	.688		
EQ58	.680		
EQ26	.658		
EQ41	.633		
EQ19	.583		
EQ36	.559	.315	
EQ1	.505		.315
EQ32		.675	
EQ59		.658	
EQ42		.593	
EQ21		.528	
EQ48		.508	
EQ6		.497	
EQ27		.473	
EQ50		.466	
EQ43	.442	.452	
EQ22	.322	.385	
EQ29		.333	
EQ8			.771
EQ35			.768
EQ12			.619
EQ14			.575
EQ4			.538
EQ57			.398

Validity - In an attempt to further examine the construct validity of the model, the correlation between factors was examined. Factors 1 and 2 correlated significantly (n = 171, r = .497, p = .0001) as did factors 1 and 3 (n = 171, r = .254, p = .001) and 2 and 3 (r = .209 p = .006). These associations were as expected; however, the coefficients are not so high as to preclude discriminant validity.

Each factor was correlated with the IRI scores from Study 1 in order to examine concurrent validity. Factor scores were used, these being a more accurate index of a person's score on a particular factor. Factor 2, 'emotional reactivity', showed a significant association with the 'empathic concern' subscale of the IRI (n = 28, r = .583, p = .001) and with 'perspective taking' (n = 28, r = .442, p =.019), but not 'personal distress' (tests are two-tailed). However, factor 3 'social skills', displayed a weak but non-significant relationship with perspective taking ($n = 28$, $r = .263$, $p > 05$). Factor 1, though did not correlate significantly with any of the IRI subscales.

A correlation matrix was also generated for total score on the Eyes task and the 3 factor scores. There was a modest but significant correlation between performance on the Eyes Test and factor 3 'social skills' ($n = 53$, $r = .273$, $p = .048$). The Eyes score, factor scores and demographics (see study 1) were also entered into a multiple regression analysis. Again, verbal IQ was the only significant predictor of performance on this task, with gender approaching significance (see Study 1 for statistics). As a further check on validity, each factor score was also correlated with predicted verbal IQ as detailed in Study 1. None of the factors showed a significant association.

Gender Differences: A 3 x 2 repeated measures analysis of variance (ANOVA) was also conducted to examine gender differences on each factor. Again there was a main effect of gender ($F_{(1, 169)} = 19.46$, $p = .001$), and an interaction between gender and the scores on the different factors ($F_{(2, 338)} = 5.85$, $p = .003$). T-tests revealed significant gender differences on 'cognitive empathy' ($t = -3.083$, $df = 169$, $p =.002$) and on 'emotional reactivity' ($t = - 4.725$, $df = 169$, $p = .001$), but not on 'social skills' ($t = 0.206$, $df = 169$, $p > .05$), with women having higher scores.

2.4 Discussion

The aim of this study was to examine the reliability, validity and factor structure of the EQ. The mean score on the EQ, in all the studies reported, was very similar to those found in the original study. Gender differences were also found for total EQ scores in mirroring the original study. These results also suggest that the questionnaire is reliable and useful across different samples and confirm that the EQ has high test-retest reliability over 10-12 months, again supporting the original findings.

The EQ showed strong concurrent validity, as evident from its moderate correlations with the previously validated 'empathic concern' and 'perspective-taking' subscales of the IRI (Davis, 1980). Furthermore, all the items entered in the final analysis distinguished between empathy and social desirability as measured by the SDS (Crowne and Marlowe, 1960), suggesting high discriminant validity. However, if the dropped items are to be used in subsequent studies, then it is important to ensure that social desirability is also measured.

The association between the Eyes Test and EQ goes some way in establishing the EQ's validity. A low but significant correlation was observed between performance on the Eyes task and total EQ score, and a weak association was also found with 'social skills'. Given that the Eyes Test includes many complex mental states and hence taps a form of cognitive empathy, this association is important. To find a relationship between the EQ, an explicit self-report measure, and the Eyes task, an implicit measure which also taps a different modality, is therefore of value.

However, the fact that EQ scores in both Study 1 and Study 4 were not significant predictors of performance on this task in multiple regression analyses reminds us of the need to control for extraneous variables. Verbal IQ as estimated by the NART was the only significant predictor of performance on this task in both of these analyses. However, the EQ score was approaching significance as a predictor in Study 1, and the lack of inter-correlation between verbal IQ and the EQ indicates that these variables are orthogonal. Although it is not possible to conclude from these data alone whether the EQ and Eyes Test are indeed tapping different constructs, the significant correlations observed hint that the Eyes task may tap *both* verbal IQ and empathy. It is also possible that the role of verbal IQ in the Eyes Test was confounded by sample selection; although this is unlikely because the occupational differences in the sample were not, in themselves, a predictor of performance. It therefore seems that both verbal IQ and total EQ scores account for some of the variance in the Eyes Task.

An exploratory factor analysis was carried out in order to elicit the underlying factor structure. The final solution accounted for a moderate amount of the total variance i.e. 41%. The first factor, labelled 'cognitive empathy', includes items that measure the appreciation of affective states e.g. 'I can tell if someone is masking their true emotion', epistemic states e.g. 'I find it easy to put myself in somebody else's shoes' and desire-based states e.g. 'I can easily work out what another person might want to talk about'. This is in line with the broader definition of 'theory of mind' as including the attribution of all types of mental state. However, it is also of interest that 'affective state' items had stronger loadings on this factor. This may also explain why no association was found between this factor and the 'perspective-taking' subscale of the IRI, as the latter is geared more towards epistemic states. Whether or not different types of mental state attribution share the same processes is an issue currently under debate (see chapter 1 and subsequent chapters).

The second factor was labelled 'emotional reactivity' as the items loading onto this factor reflect the tendency to have an emotional reaction in response to another's mental state e.g. 'seeing people cry doesn't really upset me', or 'I tend to get emotionally involved with a friend's problems'. However, items such as 'people often say I am insensitive...' and 'I can't always see why someone felt offended' also load onto this factor, illustrating how sometimes 'cognitive' empathy alone may not be enough to fully comprehend a social situation. It is possible that a full appreciation of some interactions is dependent on a certain amount of emotional resonance (see Adolphs *et al.*, 2000).

The absence of a control for 'personal distress' (Davis, 1980) in the EQ prevents us from naming factor 2 simply 'affective' empathy. From this data alone, we cannot be sure that the emotional reactions tapped in this scale are other- rather than self-orientated. One way round this potential confound is to administer the scale in conjunction with the 'personal distress' items on the IRI (Davis, 1980). This combination would give an accurate profile of empathic response. Interestingly, although this factor moderately correlated with 'empathic concern' and 'perspective-taking' on the IRI, it was not associated with 'personal distress' items. This suggests it may be tapping pure empathy after all.

The emergence of a factor tapping 'social skills' is of interest. A problem in grasping and employing everyday social skills, e.g. 'I often find it difficult to judge whether something is rude or polite', seems to be indicative of a lack of intuitive social understanding. These skills seem to require a certain amount of cognitive empathy, which is highlighted by the weak relationship between this factor and the perspective-taking subscale of the IRI. They may also rely on a knowledge base of 'social rules' required to navigate social situations. However, over-reliance on such rules, as measured in some items e.g. 'I consciously work out the rules of social situations', may be indicative of a problem with the spontaneous employment of empathy. So it seems that although an intuitive grasp of social skills is reliant on empathy and is therefore an indicator of empathic ability, an over-reliance on social rules may conversely predict a lack of empathy. It is of interest that the usual, perhaps innate, female advantage was not observed on this factor.

This attempt at reducing the items in the EQ to a few simple factors has been successful, and this series of studies confirms that the EQ provides a reliable and valid way of measuring empathy via self-report in both healthy and clinical populations. It is encouraging that Factor 1 and 2 map onto the traditional ideas of empathy, and that factor 3 provides a measure of applied empathy. Gender differences (female superiority) were also found on both cognitive empathy and emotional reactivity, supporting the previous literature. The DPD group did not differ on EQ scores as a whole, indicating an absence of a global empathy deficit. But there may well be differences on the specific components of empathy – and this will be fully examined in chapter 3.

One limitation of this study is the use of ordinal rather than continuous data in the principal components analysis, which could have resulted in a degraded solution. However, the factors that came through were easily and meaningfully interpretable. That said it should be borne in mind that this analysis is exploratory and more work needs to be done to ascertain the predictive validity of this scale. A further limitation is the rather disparate and incompletely characterised samples used including the fact that the sample in study 1 displayed a high verbal IQ as estimated from the NART. The consistency observed across studies, however, suggests that the EQ is robust to such demographic factors.

3. Conclusions

Of all the self-report measures of empathy reviewed, the EQ fits contemporary definitions of empathy most closely and is the most appropriate for clinical research. But it is important to attempt to tease out the different kinds of emotional reactivity (i.e. affective empathy vs. personal distress) and hence distinguish between empathic and other types of emotional responses to others' mental states. Unfortunately, the EQ doesn't control for this potentially confounding variable, nor does it control for social desirability. It therefore seems that the best combination of self-report measures would be to administer the EQ along with the 'personal distress' subscale of the IRI (Davis, 1980) and the SDS (Crowne and Marlowe, 1960).

Characteristics of Clinical Groups

1. Group characteristics

Contents at a glance

The data in this thesis were collected from two separate clinical groups: people with Asperger's Syndrome (AS) and those with Depersonalisation Disorder (DPD). There were 16 people in each clinical group and a control group comprising 53 healthy volunteers. Some of the data collected from the control group have already been reported in Section 2.1 of Chapter 2.

1.1 Control Group

Volunteers were recruited via circular email at the Institute of Psychiatry. In addition a database of volunteers maintained by staff at the Institute was consulted and postcards put in local newsagents, the local library and job centre. There were 53 participants of whom 29 (57%) were recruited from academic or clinical staff at the IOP and the remaining 24 (43%) from the local area and non-mental health workers at the IOP. Sixty-five percent of the women recruited were mental health workers, as opposed to 48% of the men.

The sample consisted of 25 (47%) men and 28 (53%) women. The mean age was 34.1 years (± 10.9). Forty seven (88%) stated their nationality as British, 1 (2%) Sri Lankan, 2 (4%) American, 1 (2%) Irish, 1 (2%) German and 1 (2%) Indian. 40

(77%) felt their ethnic background was white European, with 12% indicating that their ethnicity differed from their nationality, being 2 (4%) Afro-Carribean, 1 (2%) black, and 2 (4%) mixed race.

Six (11%) of the group were educated through to secondary school, 5 (10%) to further education, 14 (27%) to degree standard and 27 (52%) had a postgraduate qualification. 24 (45%) people were single, 10 (19%) co-habiting, 11 (21%) married and 4 (8%) people divorced, the remaining people either separated, widowed or with a non-cohabiting partner.

1.2 Diagnosis

Both those with Asperger's Syndrome (AS) and Depersonalisation Disorder (DPD) were diagnosed by psychiatrists using standardised ICD-10 criteria or DSM IV criteria. In addition, valid and reliable 'gold standard' measures were also administered.

1.2.1 Asperger's group

Diagnosis

Asperger's Syndrome is often described as a mild autistic spectrum disorder, because it mirrors more severe forms of autism except that diagnosis is given on the basis of normal language development. All the volunteers were diagnosed with Asperger's Syndrome (AS) by a psychiatrist using the following diagnostic criteria (ICD-10):

ICD-10 (F8.45)

Diagnosis is based on the combination of a lack of any clinically significant general delay in language or cognitive development plus, as with autism, the presence of qualitative deficiencies in reciprocal social interaction and restricted, repetitive, stereotyped patterns of behaviour, interests, and activities. There may or may not be problems in communication similar to those associated with autism, but significant language retardation would rule out the diagnosis.

Demographic details

The mean age of this group was 37.7 years (±10.1); there were 14 men (87.5%) and 2 women (12.5%). All the participants in this group were British, and they all indicated that their ethnicity would fall into the category of white European. Four were unemployed (25%). They further numbered 2 civil servants (12.5%), 2 shop workers (12.5%), 2 students (12.5%), 1 charity worker (6.3%), 1 factory worker (6.3%), 1 gardener (6.3%), 1 admin worker (6.3%), 1 family support worker (6.3%) and 1 retired person (6.3%). Three (18.8%) of the participants had been educated through to secondary school, 5 to further education (31.3%), 6 (37.5%) to degree level and 1 (6.3%) had a postgraduate qualification; there was 1 instance of missing data. Thirteen (81.3%) described their relationship status as single, 1 (6.3%) as having a non-cohabiting partner and 2 (12.5%) were married.

Clinical Characteristics

Six people were also administered the Autism Diagnostic Interview-Revised (Lord *et al.*, 1994), an interview for caregivers of individuals with possible pervasive developmental disorder and 3 were assessed using the Autism Diagnostic Observation Schedule (Lord *et al.*, 1989) – a tool designed to allow clinicians to record their observations in order to reach a conclusion about diagnosis. These are both accepted 'gold standard' measures of autistic spectrum disorders (ASDs), the main difference being that the former utilises reports from caregivers and is therefore particularly useful for younger people, whereas the latter is often used when no caregiver information is available. All met the criteria for AS using these measures. Informant information was not available for the remaining 7 people.

1.2.2 Depersonalisation group

Diagnosis

Depersonalisation Disorder (DPD) is severe depersonalisation associated with functional impairment. Depersonalisation as a symptom can be both primary and secondary to other psychiatric illness (Lambert *et al.*, 2001a). It can occur in healthy individuals after taking illicit drugs (Medford *et al.*, 2003), and during times of extreme stress such as 'near death' experiences (Noyes and Kletti, 1977). When depersonalisation is severe and the primary symptom, a diagnosis of DPD is given, which is classified as one of the dissociative disorders in the DSM IV but as a neurotic condition in the ICD-10.

All participants in this group were diagnosed with DPD following a full psychiatric interview and assessment (by Dr Sierra or Dr Medford at the Depersonalisation Research Unit at the Maudsley Hospital, London) at the request of the individual and their referring clinician. A final clinical diagnosis was made according to DSM-IV (1994) criteria 300.6 and ICD-10 using the Present State Exam or PSE (DSM-IV, 1994). The PSE includes items for depersonalisation and derealisation: 0 = not present, 1 = moderately intense or transient, and 2 = intense and persistent. Our case definition required a total score of 2 or above without an obvious additional clinical diagnosis or prominent non-dissociative symptomatology.

Demographic details

The DPD group consisted of 16 people including 4 (25%) women and 12 men (75%). The mean age of the group was 32 years (± 7.8). There were 14 (81.3%) people of British nationality, 1 (6.3%) German and 1 (6.3%) Australian. 11 (68.8%) people were single, 2 (12.5%) were co-habiting, 2 (12.5%) married and 1 (6.3%) was separated. 2 (12.5%) participants were educated to secondary level, 2 (12.5%) had completed further education, 9 (56.3%) had been educated to degree level and 3 (18.8%) had postgraduate qualifications. There were 3 (18.8%) students, 3 (18.8%) unemployed and a photographer, a graphic designer, an education welfare officer, a news analyst, a sales manager, a gardener, a nursery assistant, a company director, a technical writer, and a housing worker.

Clinical Characteristics

All participants (including those in the larger DPD group described in Chapter 2 Section 2.4) were given the Dissociative Experiences Scale version II (DES - Bernstein and Putnam, 1986, Carlson and Putnam, 1993), the Cambridge State and Trait Depersonalisation scales (Carlson and Putnam, 1993), the Beck Anxiety Inventory (BAI - Beck *et al.*, 1988a) and Beck Depression Inventory (BDI - Beck *et al.*, 1988b). These were sent to participants prior to referral to the clinic, and the information was entered into a comprehensive database maintained by Dr Dawn Baker at the Depersonalisation Research Unit, Institute of Psychiatry, London. Some of the data in this database have been reported elsewhere (Lambert *et al.*, 2001a, Lambert *et al.*, 2001b). The DES, BAI and BDI scores for the larger DPD group can be found in section 2.4. of chapter 2

The mean DES scores for the smaller DPD group described here were 23.9 (± 13.9), with the mean DES taxon score being 24.9 (± 16.6) and the DES DP/DR factor being the highest at 36.4 (± 22.1). There were 5 instances of missing data resulting from people failing to complete and return the measure. Background and normative data for the DES scale can be found in section 2.4 of chapter 2 – but to summarise: an overall score above 30 is indicative of severe dissociation with the taxon having a cut-off of 13 for dissociative disorders, a score above 8 on the DP/DR scale being suggestive of a dissociative disorder. This group, therefore, tended to score in the dissociative disorders range for the taxon and the DP/DR factor, but not on the DES overall. This is consistent with previous data reported for the larger DPD group and other samples (Lambert *et al.*, 2001a, Lambert *et al.*, 2001b).

The Cambridge Depersonalisation 'trait' and state' self-report questionnaires are devised to tap depersonalisation independently from other dissociative disorders. The trait scale measures the frequency and duration of DPD symptoms of the last 6 months, whilst the state scale simply measures the frequency of symptoms over the last week. The median trait score in the original paper was 113 and the cut-off point for depersonalisation is 70. The median score in this sample was 130, therefore well within the depersonalisation range. There were 2 instances of missing data.

Lastly, participants were also given the BAI and BDI, due to the co-morbidity between DPD, depression and anxiety (Baker *et al.*, 2003, Lambert *et al.*, 2001a). A score below 11 on either scale is considered within the 'normal' range, whereas a score above 30 is classed as 'severe'. The mean score on the BAI was 22 (± 9) and for the BDI it was 22.8 (± 10.7). Again, there were 2 cases of missing data. This therefore suggests that this group suffered mild to moderate depression and anxiety, mirroring the characteristics of the larger cohort reported elsewhere (Lambert *et al.*, 2001b).

2. Self-report measures

Following up the conclusions of Chapter 2, the EQ, the Social Desirability Scale and the 'personal distress' items of the IRI were given to both clinical groups. These were administered as part of a larger testing session, and participants either completed them during the testing phase or took them home to fill in and return.

2.1 Crowne and Marlow social desirability scale

Background

The background to this scale has been reviewed in Chapter 2. This measure was administered in order to examine socially-desirable responses in both of the clinical groups. It is important to confirm that all groups are within normal range on this variable, and that there are no group differences which might confound comparisons.

Analysis

There was 1 case of missing data so the following results are based on a sample size of n = 15. The mean scores were 11.71 (± 6.3) for the DPD group, 13.57 (± 5.1) for the control group, and 14.44 (± 6.3) for those with Asperger's Syndrome. A one-way ANOVA was also performed to examine group differences on this measure. The overall F test was not significant ($F_{(2, 82)} = .786$, $p > .05$).

Discussion

The tendency to respond to self-report measures in a socially desirable way is within the normal range for each of the clinical groups. Neither clinical group's scores were significantly different from the control group on this measure, confirming the validity of comparisons between these groups on self-report measures.

2.2 The empathy quotient

Background

For an overview of the background, reliability and validity of this measure see Chapter 2. The prediction is that the AS group will score significantly lower overall, also on factor 1 'cognitive empathy', and factor 3 'social skills'. But it is not easy to predict how the AS group will perform on the 'emotional reactivity' items, as the evidence is not conclusive (see chapter 1). However, many of the items also include a 'theory of mind' component, so it is likely that this group will also perform significantly worse on this factor too. It is expected that the DPD group may also have a significantly lower score on the 'emotional reactivity' items due to the symptom of emotional blunting they frequently report.

Analysis

The total EQ scores for the controls, the AS group and the DPD group can be found in Table 3.1. The results were stratified by gender, as a female advantage has been found consistently on this measure (Baron-Cohen and Wheelwright, 2004 in press, Baron-Cohen *et al.*, 2003b).

Table 3.1 Mean and SD total EQ scores for controls and clinical groups stratified by gender

		gender			
		male		female	
		total score on the EQ		total score on the EQ	
		Mean	SD	Mean	SD
group	Control	41.3	10.1	50.6	9.2
	DPD	41.4	8.8	43	5.6
	AS	26.4	12.6	36	17.0

A univariate ANOVA with both group and gender as fixed independent variables and total EQ score as the dependant variable, showed significant differences between groups ($F_{(2, 78)} = 6.79$, p=.002) with *post hoc* Scheffé tests revealing that the AS group scored significantly lower than both the control group and the DPD group. Gender was also significant, with women scoring consistently higher than men ($F_{(1, 78)} = 4.20$, p = .044). However, there was no significant interaction between group and gender, confirming that the apparent difference in scores between the female control and DPD groups was not significant (NB: there were 2 women in the AS group).

Based on the results of the data screening and factor analysis in Chapter 2, mean scores for each EQ factor were also calculated for each group (see Table 3.2).

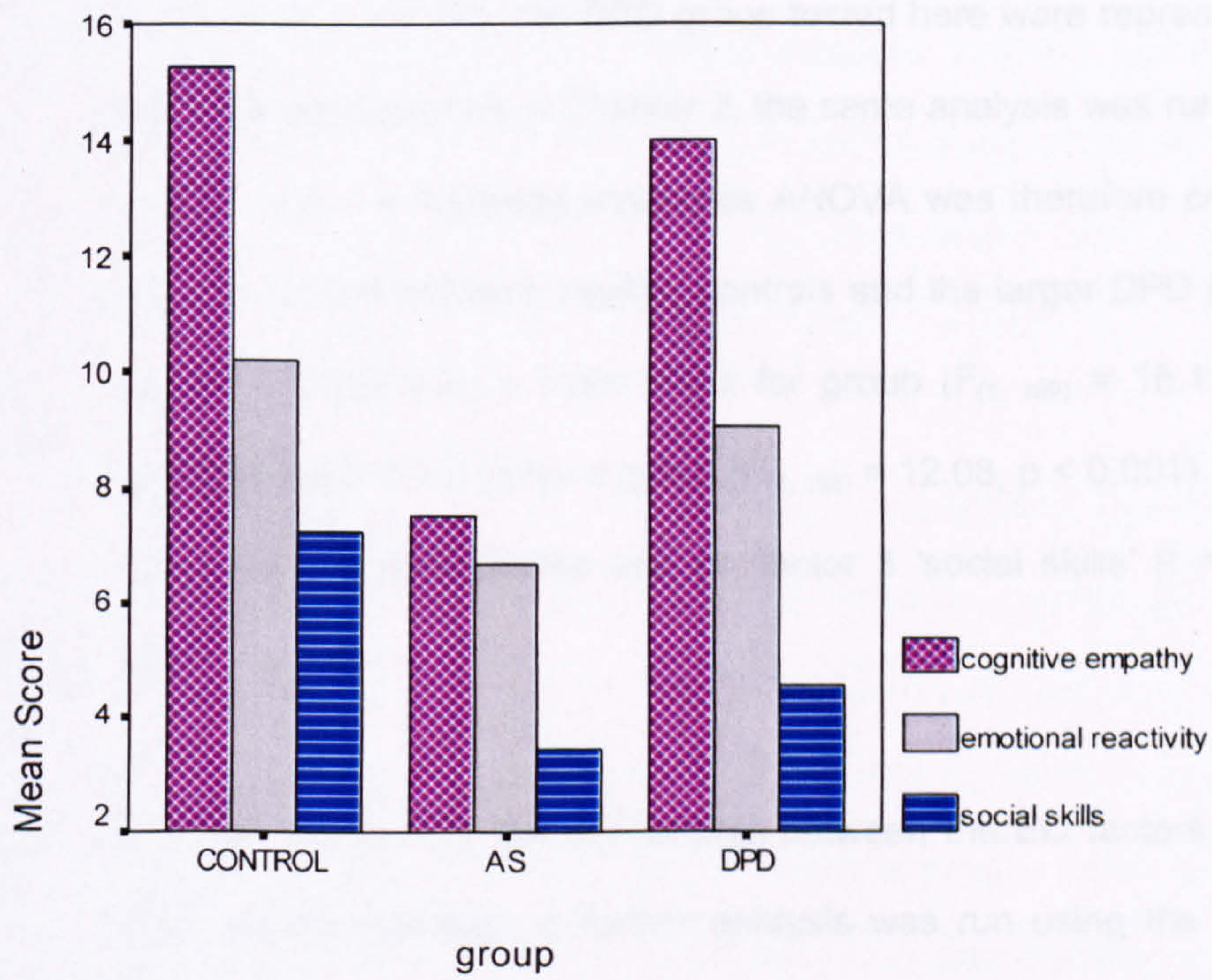
Table 3.2 Mean and SD EQ factor scores for each group

		factor 1 - cognitive empathy		factor 2 - emotional reactivity		factor 3 - social skills	
		Mean	SD	Mean	SD	Mean	SD
group	control	15.2	4.8	10.2	4	7.2	2.6
	DPD	14.9	5.5	9.5	3.4	4.3	2.5
	AS	7.5	7.1	6.7	3.4	3.4	3.4

A 3 x 3 repeated measures ANOVA was used to ascertain whether there were any significant group differences on the three EQ factor scores, with gender entered as a co-variate. As expected, there was a main effect for group ($F_{(2, 80)} = 11.77$, $p = .002$) and a significant interaction between EQ factors and group ($F_{(4, 160)} = 3.95$, $p = .004$). The interaction between EQ factors and gender was also significant ($F_{(2, 160)} = 4.66$, $p = .011$).

In order to tease out the simple effects, a one-way ANOVA with *post hoc* Scheffé tests ($p = .05$) was conducted for group by factor. The AS group differed significantly from both the other groups on all factors, and the DPD group differed significantly from the control group on factor 3 'social skills' (see Figure 3.1).

Figure 3.1: Mean scores for each group on each EQ factor



T-tests also revealed significant differences between males and females on 'cognitive empathy' ($t = -4.29$, df 82, $p < .001$ - values for unequal variances), 'emotional reactivity' ($t = -.3.03$, df 82, $p = 0.003$), and 'social skills' ($t = -.1.98$, df 82, $p = .051$).

An item-by-item analysis of the specific 'emotional reactivity' items revealed that people with DPD did not perform significantly worse than controls on any one of them. However, the AS group scored significantly lower than controls on the following items:

- I get upset if I see people suffering on news programmes. ($t = 2.33$, df 67, $p = .023$)
- It is hard for me to see why some things upset people so much. ($t = 3.1$, df 67, $p = .003$)
- Other people often say that I am insensitive...I don't always see why.($t = 3.19$, df 67, $p = .002$)
- I can't always see why someone ....felt offended by a remark. ($t = 3.48$, df 67, $p = .001$).

In order to check that the DPD group tested here were representative of the larger DPD sample reported in Chapter 2, the same analysis was run with the larger DPD group. A 3 x 2 repeated measures ANOVA was therefore conducted to examine the differences between healthy controls and the larger DPD group on the 3 factor scores. There was a main effect for group ($F_{(1, 169)} = 15.11$, $p = 0.001$) and a significant effect for factor x group ($F_{(2, 338)} = 12.08$, $p < 0.001$). T-tests showed that the main group difference was on factor 3 'social skills' ($t = 6.66$, df = 169 $p = 0.001$).

In order to examine the relationship between the EQ factors and other indices of DPD symptomatology, a further analysis was run using the data from the larger DPD group. A correlation matrix was generated for the DES scores and each EQ factor. There were significant correlations between 'social skills' and DES mean ($n = 52$, $r = -.348$, $p = .011$), amnesia ($n = 52$, $r = -.291$, $p = .036$) and 'absorption' ($n =$

52, r = -.369, p =.007). The relationship between the BDI, BAI and each factor was examined. This revealed that 'emotional reactivity' showed a significant positive correlation with anxiety scores (n = 52, r = .313, p = 0.024) and 'social skills' showed a significant negative correlation with depression scores (n = 45, r = - .346, p = 0.012).

Discussion

As predicted, the AS group scored significantly lower on both the total EQ score and each of the individual EQ factors. The DPD group did not score any lower than the control group on total EQ score, mirroring the findings from the larger DPD group in Chapter 2. This suggests that people reporting symptoms of depersonalisation who suffer mild to moderate levels of anxiety and depression do not suffer a global empathy deficit. In fact, they show the same pattern of results on the EQ as the psychologically healthy population detailed – including gender differences. The apparent difference between total EQ scores for the female DPD and control groups was not significant and probably arose as an artefact of sample selection because the female controls were drawn from a group that may have been particularly empathic i.e. mental health workers.

The AS group performed according to predictions reporting significantly less proficiency on items tapping 'cognitive empathy', 'emotional reactivity' and 'social skills'. This difference was most prominent with 'cognitive empathy' and social skills'. This is in line with the existing literature, which suggests that people with ASDs have a deficit in employing a 'theory of mind' – particularly as social skills are constantly reliant on this mechanism. An item-by-item analysis of 'emotional reactivity' revealed that the AS group scored significantly lower on items that require the appreciation of the representational qualities of other peoples' minds, as expected. However, their performance on items that are not so reliant on such an ability was not significantly different. Here are some examples:

- Seeing people cry doesn't really upset me. *Reverse scored*
- I tend to get emotionally involved with a friend's problems.
- I really enjoy caring for other people.
- I usually stay emotionally detached ...watching a film. *Reverse scored*

This may account for the fact that, although significant, the AS groups score was only 1 SD below the control group on this factor, a smaller difference than that observed on the other 2 factors.

Contrary to predictions, there were no group differences on 'emotional reactivity' in the DPD group and the item-by-item analysis did not shed any light on this. There were, however, significant differences on 'social skills'. The relationship between DPD and indices of empathy is particularly hard to disentangle when one considers the co-morbidity of both depression and anxiety in DPD. But the BDI scores had a negative association with 'social skills' and it is therefore possible that depressed mood accounts for some of the group differences on 'social skills', perhaps due to motivational factors.

Alternatively, the DES mean and DES absorption scores also showed a negative association with 'social skills', suggesting that there may be some specific components of DPD that lead to people to report a problem in this area. However, no firm conclusions can be drawn about either explanation until further studies are conducted using *objective* measures of empathy. This is necessary to overcome the potential criticism of all self-report measures, which is that they rely on biased subjective evaluations. One way of overcoming this problem would be to obtain objective reports from someone who knows the respondent well.

The gender differences observed in previous studies were replicated, confirming that this sample is representative of the normal population. It was interesting to see female superiority in both 'cognitive empathy' and 'emotional reactivity', indicating that the gender differences ordinarily observed do not result solely from an advantage in one domain. This is consistent with models that see the male and female brain as varying on a continuum from 'systemising' through to 'empathising' (see section 3.3 below).

Finally, the BAI scores showed a positive relationship with 'emotional reactivity' – which is important when one considers that this factor may also include a tendency to experience 'personal distress' as a result of others' emotional states. This finding emphasises the need to control for anxiety when attempting to measure 'affective empathy'. More attention needs to be paid to the effects of both anxiety and depression on empathy, and this may be of importance in clinical settings.

2.3 Interpersonal reactivity index

Background

The literature reviewed and findings from Chapter 2 suggest that it is necessary to administer the 'personal distress' items of this scale along with the EQ. This allows conclusions to be drawn as to the nature of factor 2, which is tentatively labelled 'emotional reactivity', as opposed to 'affective empathy', due to its failure to distinguish between self- and other- orientated reactivity. It remains unclear how the AS group will score on this measure thus this analysis is exploratory and the predictions are two-tailed. However, it is expected that the DPD group will score lower than the control group on this measure, as the emotional blunting reported by this group is thought to extend to negative emotions about the self.

Measures

Five items from the IRI were chosen due to time constraints. These were:

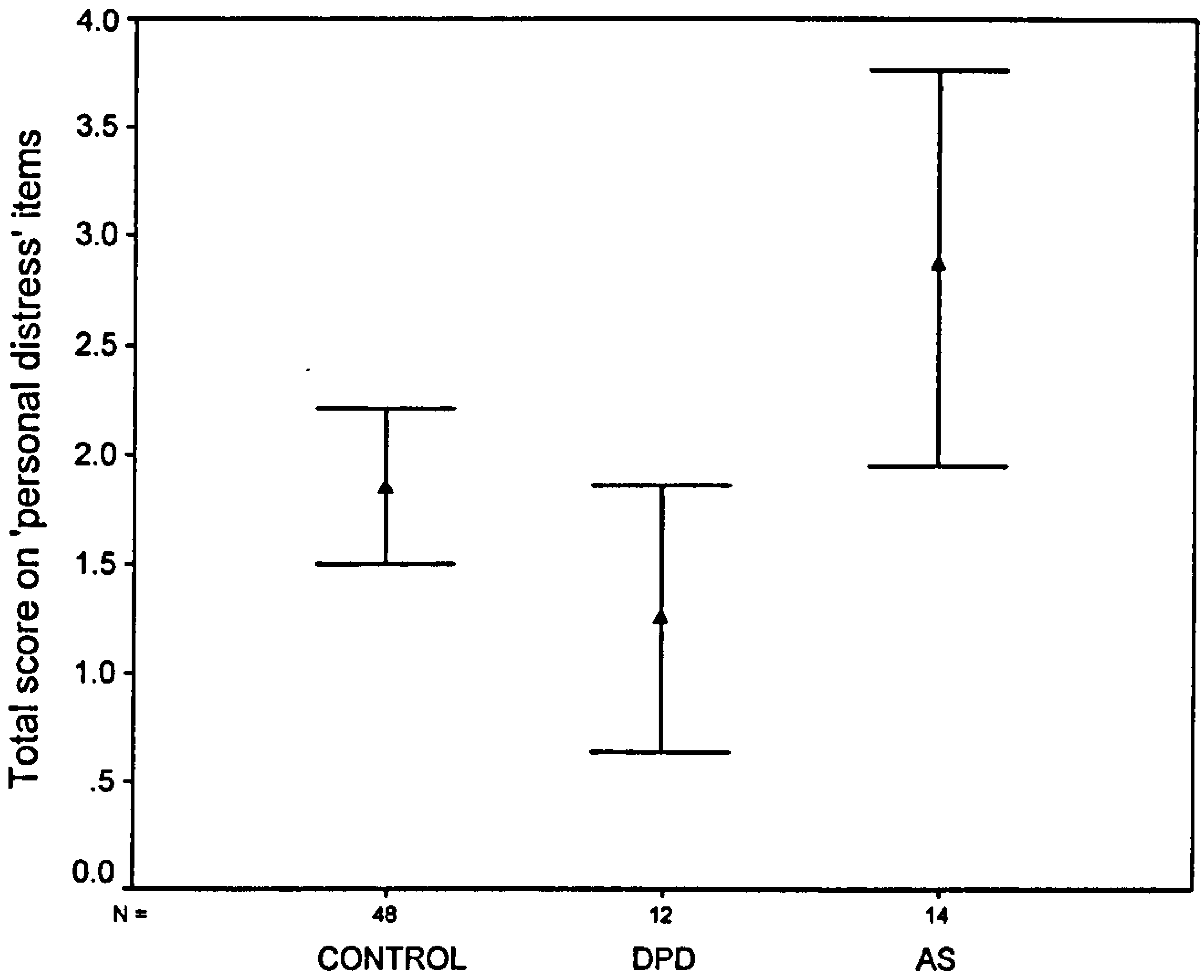
- Do you feel apprehensive and ill at ease in emergency situations?
- When you see someone get hurt do you tend to remain calm? *reversed*
- Does being in a tense and emotional situation scare you?
- Do you feel helpless when you are in the middle of a very emotional situation?
- Are you usually pretty effective at dealing with emergencies? *reversed*

These items were dispersed among questions concerning the respondents' attitudes' and opinions. Participants were required to endorse a simple 'yes' or 'no' to the items, as opposed to a Likert scale – again due to time restraints. Four people from the DPD group and 2 from the AS group failed to return the questionnaire. Only 48 control participants were given the IRI, as opposed to 53, as the first 5 respondents formed part of a pilot study which omitted this measure.

Analysis

The mean and SD scores on this scale are shown in Figure 3.2. A one-way ANOVA was performed with *post hoc* Scheffé tests to examine the simple effects. There was a main effect for group ($F_{(2, 73)} = 5.58$, $p = .006$) and the differences were mainly between the AS and control group, with the former scoring significantly higher on this index.

Figure 3.2: Mean scores and 95% confidence intervals for 'personal distress' items



Discussion

People with AS score significantly higher on items that measure the propensity to react to negative situations of others with 'personal distress', i.e. a self-orientated reaction. Contrary to predictions, people with DPD scored slightly, but not significantly, lower on this measure. This self-orientated response is incompatible with empathic responding (although it is possible that it shares some of the same mechanisms) and these findings therefore suggest that the AS group may have difficulty responding empathically to others' emotional states. However, it must be borne in mind that some of the questions may tap the ability to employ a representational 'theory of mind' (ToM) which may also account for the findings in the AS group.

The fact that the DPD groups scores were in the predicted direction but of small effect may suggest that it is possible that the small number of items and yes/no format lacked sensitivity. It may therefore be worth replicating this finding using the original measure, i.e. 7 items and a Likert scale. If this trend is indeed due to group rather than random differences, one explanation would be that people with DPD have a problem generating any emotional reaction to others' distressing states, whether self- or other-orientated.

3. Cognitive function

In order to ensure any experimental effects were not due to global deficits, measures of IQ, conditional reasoning and physical problem-solving were administered to both clinical groups. Although all the novel tasks have their own internal control (see subsequent chapters), it was considered important to ensure that cognitive functioning as a whole was intact before reaching any conclusions about emotional reasoning in each of the clinical groups.

3.1 Verbal intelligence quotient

Both clinical groups and the control group were given the National Adult Reading Test (Nelson, 1982) as part of the experimental session. This measure yields an estimate of verbal IQ. A measure of verbal IQ was deemed necessary, as much of the material used in the study was in written form.

This NART was always given early on in the testing session, after about 10 minutes had been spent clarifying demographic details, obtaining consent and performing one short experimental task. Participants were asked to read aloud 50 irregular sounding words e.g. 'ache'. The experimenter marked the score sheet as the participant proceeded through the list. The participant was also recorded onto a Sony Mini Disc player to enable the experimenter to re-score the test if necessary. The mean and SD Verbal IQ for each group can be found in Table 3.3.

Table 3.3 Mean scores for verbal IQ predicted from national adult reading test

		predicted verbal iq from NART	
		Mean	SD
group	control	121.1	5.1
	dp	117.8	5.2
	as	117.4	6.5

The data show that all the groups are above the average range for IQ of 100 (±15). As all the groups have similar IQ's and are well-matched on this variable, comparison between groups on different variables will be fruitful.

3.2 Conditional reasoning tasks

The use of reasoning tasks to study certain aspects of mental disorder has burgeoned over the last few years. This is in part due to the increase in cognitive models of mental illness, i.e. depression, anxiety, delusional states (Nelson, 1982, Dalgleish and Power, 1999). Conditional reasoning tasks have also been used to this end. Such tasks lead respondents to make systematic errors, which are reliably reduced and exacerbated under different conditions.

Kemp, Chua, McKenna & David (1997) used a conditional reasoning task to examine the reasoning style of people with delusional beliefs. There are four forms of conditional reasoning statement, with two leading to valid inferences and two inducing the endorsement of fallacies. The two valid forms are:

- Modus Ponens/Form A (if p then q)
- Modus Tollens/Form not B (If not q then not p).

And the invalid forms are:

- Denying the Antecedent/Form not A (if not p then not q)

- Affirming the Consequent/Form B (If q then p).

The authors manipulated the tasks to include both additional and alternative premises and statements with emotional content. The provision of alternative premises (if p or r then q) usually reduces the endorsement of fallacies but has no effect on the valid forms. On the other hand, the provision of additional premises (if p and r then q) normally suppresses valid inferences. Emotional content also leads to inaccurate responses in both healthy volunteers and clinical groups.

An example of an alternative statement is:

"if it rains I will get wet, if it snows I will get wet. It does not rain, therefore...we can't say'.

The alternative premise of 'if it snows...' highlights the fallacy by showing people that there are other options – this therefore reduces the amount of fallacies endorsed.

Statements with additional premises, on the other hand, reduce the amount of valid inferences, as they highlight alternative possibilities leading respondents to deviate from the usual reasoning pattern. An example of an additional statement is:

"if she meets her friend she will go to the play, if she has enough money she will go the play. She meets her friend ... therefore she goes to the play"

In the Kemp *et al.,* (1997) study, the clinical group, i.e. people suffering from delusions, were found to make significantly more errors than a control group both overall and on the emotional statements. This, along with the data from other reasoning tasks in this study, was interpreted as possible evidence for a cognitive bias in people with delusions. However, the authors were clear to point out that all groups made errors on statements with emotional content. Healthy individuals high on schizotypal traits have also been found to display reasoning biases on such tasks (Lawrence and Peters, 2004 in press). In order to make predictions about the current clinical groups' performance, it is necessary to examine more closely the basis of the common errors on such a task. The over-endorsement of fallacies in the statements with emotional content is an ideal starting point.

In this kind of task, it seems that highly emotive material prevents participants from displaying the same accurate reasoning strategy that they previously applied to non-emotive statements. It is entirely possible that this is due to the activation of cognitive and neural mechanisms specialised for material with emotional content. There is a wealth of literature suggesting that the processing of emotional material has distinctive cognitive and neural correlates (Dolan, 2002). It is probable, however, that people with DPD will not show this pattern of errors with emotional material. If we assume that physiological arousal is the trigger for the differential processing of emotional material (it being the main quality that distinguishes this mental state from others) then emotional blunting (or lack of physiological arousal) may avert this effect. This would result in people with DPD showing an advantage on such a task.

Another error consistently shown by both controls and the clinical group in the Kemp study was the suppression of valid inferences when an additional premise was provided. It seems that additional premises remind participants of other 'real-world' possibilities that could affect the outcome of the given situation. Presumably this then leads respondents to consider the additional premise in forming their answer, despite instructions to ignore it. Participants are then tempted to deviate from their previously accurate reasoning strategy in order to give an answer that is consistent with *both* premises and hence externally coherent. The cognitive strategy needed to avoid this mistake is simply to ignore the additional statement and resist the temptation to process the premise in context.

This attribute is part of the cognitive style documented in those with AS who show what has been termed 'weak central coherence' (~~Dolan, 2002~~ Frith, 2003). Non-autistic people tend to process disparate facts globally and attempt to construct meaning between them. This cognitive strategy probably underlies the pattern of errors observed on tasks such as these. However, several studies show that people with ASDs do not share this tendency, and instead process information in a piecemeal fashion (Happe, 1999). Jolliffe & Baron-Cohen (2001) found that both people with autism and Asperger's Syndrome (split into 2 separate groups) were significantly less likely to integrate fragments of a drawing in order to recognise it despite being able to recognise the same drawing as a single part. People with autism also show an advantage on the Embedded Figures Task (Ring *et al.*, 1999) which requires the extraction of an image from a series of lines designed to disguise the object. This processing style also transfers to other domains such as verbal comprehension. Happé (1999) found that a sample of 16 people failed to process information in a context-dependent manner, leading to more inaccuracies on homographs e.g. there was a *tear* in her eye vs. there was a *tear* in her dress.

If people with ASDs do show 'weak central coherence', then we might predict that they will show an advantage on conditional reasoning statements with additional premises. If additional premises ordinarily lead participants to try and generate a coherent solution which is consistent with the additional information then people with ASDs are less likely to employ such a strategy and more able to ignore the additional information and process the premises in isolation. This should lead to enhanced performance. A similar explanation was given by Scott, Baron-Cohen, & Leslie (1999) for their finding that children with autism outperformed verbal-age matched controls on a counterfactual reasoning task. Success on these tasks requires one to suspend 'real-world' empirical knowledge and process the premises independent of information in long-term memory. It seems reasonable, therefore, that the same cognitive style that leads to success on counterfactual reasoning tasks will result in an advantage on conditional reasoning statements with additional premises.

It was hypothesised that:

- Neither clinical group will perform significantly worse than controls overall on the conditional reasoning statements.
- The DPD group will make fewer errors on the tasks with emotional content.
- The AS group will make fewer errors on tasks with additional premises.

Materials

There were 48 items, including some with emotional content and some with neutral content. Of the neutral statements, there were 3 sets of 4 in the simple form, 3 sets in the alternative form and 3 sets in the additional form. There was 1 set in the simple form with emotional content, 1 alternative set with emotional content, and lastly 1 additional set with emotional content. The statement with emotional content is as follows;

If she is raped she will go to the police.

She is raped. (MP)

- She will go the police. (T)
- She will not go to the police.
- Can't say.

The alternative version would read:

If she is raped she will go to the police.

If she is mugged she will go the police (alternative)

She is raped. (MP)

The additional version would read:

If she is raped she will go to the police.

If she is able to walk she will go the police. (additional)

She is raped. (MP)

Procedure

Participants were instructed as follows: "assuming that all the statements are true, please tell me which answer follows on from the given statement". The presence of additional and alternative statements was not explicitly marked, but was implicit from the statement. If prompted regarding additional and alternatives, the participants were told that the statements are to be taken as true in their own right.

This reasoning task was given as part of a larger testing set. For both groups, the task was given in two sections to allow participants a break, i.e. 1 - 24 followed by 25 - 48. The DPD group was administered the task before and after some emotional diary extracts (see chapter 4). The latter were counterbalanced using a 4 x 4 Latin Square – which means the conditional reasoning task was followed or preceded by 4 different extracts at an equal frequency. The AS group were given the task slightly later in the battery, after the diary extracts and the first and second set were broken up by a physical problem-solving task (see below).

Analysis

The mean scores for each group in the neutral conditions can be found in Table 4. The normative data are taken from Kemp *et al.*, (1997). To allow comparison with the control data, mean as opposed to median percentages are given, despite the fact that the percentages are derived from a rather small range of scores.

Table 3.4: Mean percentage of valid inferences and fallacies endorsed on neutral statements

	group			
	dp		as	
	Mean	SD	Mean	SD
% valid simple inferences	85.9 (89*)	15.7	95.3 (89)	12.8
% simple fallacies endorsed	72.7 (81)	27.5	64.8 (81)	44.1
% valid alternative inferences	78.1 (86)	19.0	87.5 (86)	16.7
% alternative fallacies endorsed	37.5 (39)	36.3	39.6 (39)	41.7
% valid additional inferences	68.8 (73)	25.7	84.4 (73)	18.7
% additional fallacies endorsed	50 (57)	34.4	52.1 (57)	46.3

* Figures in parentheses are mean scores from the control group in the Kemp et al., (1997) study.

The mean percentages for each condition with emotional content are given in Table 3.5, as before, mean scores are given and the control data is provided in brackets.

Table 3.5: Mean percentage of valid inferences and fallacies endorsed on emotional statements

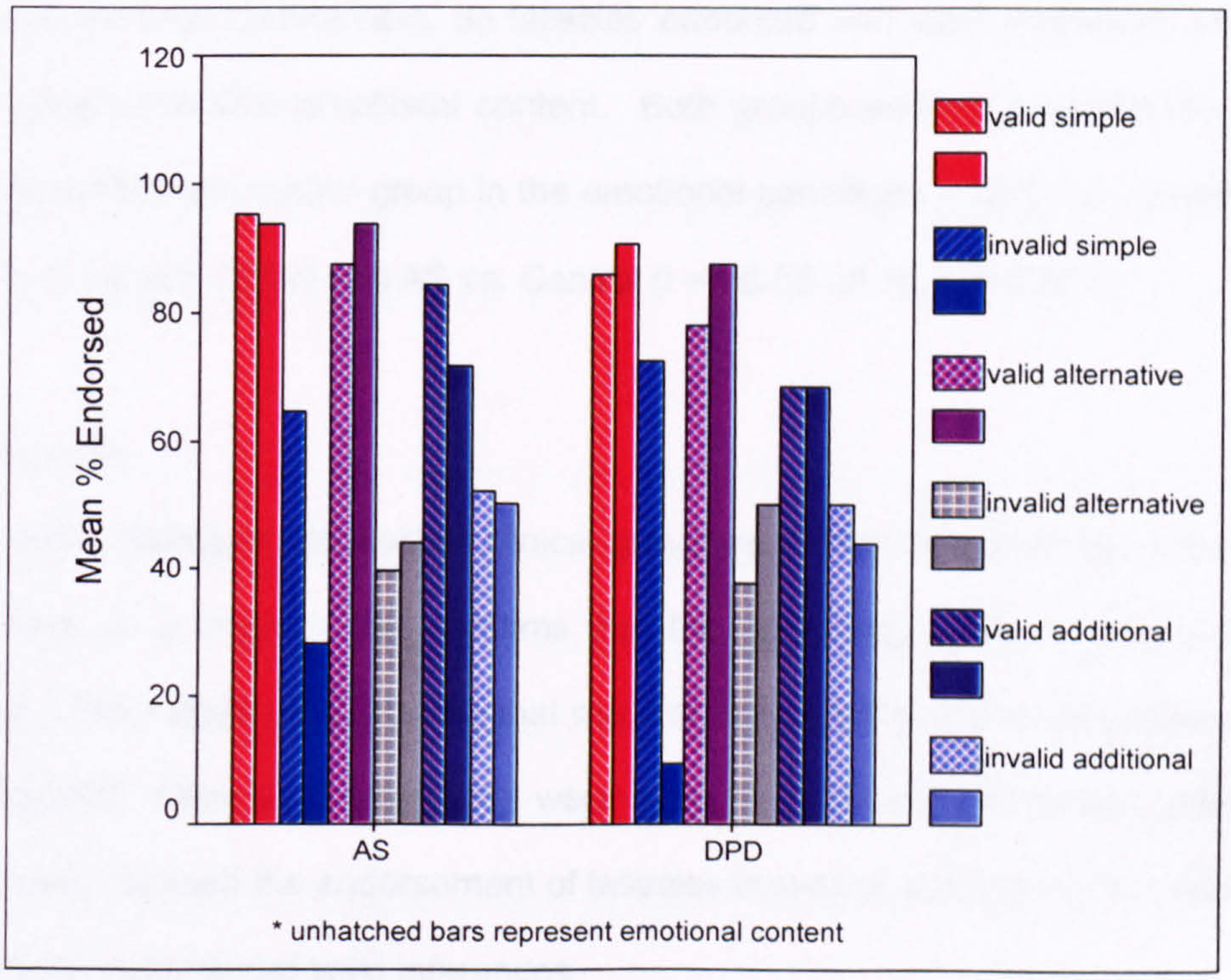
	group			
	dp		as	
	Mean	SD	Mean	SD
% valid simple inferences	90.6 (84)	20.2	93.8 (84)	17.1
% simple fallacies endorsed	9.4 (91)	27.2	28.1 (91)	44.6
% valid alternative inferences	87.5 (88)	22.4	93.78 (88)	17.1
% alternative fallacies endorsed	50 (44)	44.7	43.8 (44)	44.3
% valid additional inferences	68.8 (63)	40.3	71.89 (63)	25.6
% additional fallacies endorsed	43.75 (78)	44.3	53.1 (78)	49.9

A repeated measures ANOVA was conducted to examine the variation between the clinical groups and the success of the experimental manipulations. There were differences between the category of response, i.e. valid vs. invalid fallacious inferences ($F_{(1, 30)} = 76.26$, $p < .001$), and the means demonstrate this to be due to fewer fallacious than valid inferences overall. Whether or not the statement had emotional content was also significant ($F_{(1, 30)} = 6.349$, $p = .017$), and inspection of the means showed that *fewer* invalid fallacies were endorsed in the emotional conditions.

The interaction between category of response (i.e. valid inferences vs. invalid inferences) and the provision of extra premises, i.e. simple, additional or alternative premises, was also significant ($F_{(2, 60)} = 6.02$, $p = 0.04$). It seems that there were fewer valid inferences with additional statements than without additional statements [simple form ($t = 3.24$, df 31, $p = 0.003$); alternative form ($t = 3.5$, df 31, $p = 0.001$)]. However, the number of fallacies endorsed in the presence of additional statements did not differ from the number endorsed without additional statements [simple form (t = -.745, df 31, p >.05); alternative form (t = -1.28, df 31, p > .05)].

There was also an interaction between category and emotional content ($F_{(1, 30)}$ = 10.87, p = 0.03) and paired t-tests revealed that there were significantly more fallacies endorsed in the neutral conditions as opposed to the emotional conditions (t = - 3.268, df 31, p = .003). Lastly, there was a significant interaction between category, the provision of extra premises (simple, additional, alternative), and emotional content ($F_{(2, 60)}$ = 10.77, p = 0.001). The means suggest this to be due to the markedly lower rates of invalid fallacies endorsed by both groups in the simple emotional statements (see Figure 3.3 – invalid simple).

Figure 3.3: Mean percentage of valid and invalid inferences on conditional reasoning statements



There were no significant interactions between group and category response (F (1, 30) = .357, $p > .05$), group and emotional content ($F_{(1, 30)} = .181$, $p > .05$), and between group and extra premises ($F_{(2, 60)} = .116$, $p > .05$). In order to examine any group effects with control data, and hence the main hypotheses more closely, one sample t tests were conducted.

Firstly, the AS groups performance on valid inferences made with additional statements was examined. There was a significant difference in the neutral condition, as the AS group endorsed more valid inferences in the presence of additional premises than the controls (t = 2.43, df 31, p = .028). However, this was not the case for the DPD group (t = -.66, df 31, $p > .05$), and this finding did not hold for statements with emotional content.

One sample t-tests were also conducted to compare the data from both clinical groups with the control data, on fallacies endorsed and valid inferences made on the statements with emotional content. Both groups endorsed significantly fewer fallacies than the control group in the emotional conditions – DPD vs. Control (t = -7.11, df 15, p = .0001) and AS vs. Control (t = - 6.03, df 15, p = 0.001).

Discussion

The main finding is that neither clinical group performed less well than controls on the task as a whole. This confirms that the reasoning ability of both groups is intact. The results also indicate that most of the experimental manipulations were successful. More valid inferences were made than fallacies endorsed, alternative premises reduced the endorsement of fallacies in neutral statements and additional premises suppressed valid inferences.

The prediction regarding the AS group's performance on the statements with additional premises was also upheld. The controls but not the AS group made fewer valid inferences in the presence of additional statements; in fact, one sample t-tests showed that, with the neutral statements, people with AS made more valid inferences than controls in the presence of additional statements. This is likely to result from their tendency to process information in isolation from its context.

There was also a main effect for emotional content, and there were no differences between the groups when reasoning with emotional material. It seems that fewer fallacies were endorsed by both groups with statements that had emotional content. This was most marked with simple statements, resulting in a significant interaction effect for category of response (valid vs. fallacious) x extra premises (additional, alternative or simple) and emotion (neutral vs. emotional). One sample t-tests using the Kemp *et al.,* (1997) normative data confirmed that both clinical groups endorsed fewer fallacies than controls in the statements with emotional content. In neither clinical group was emotional content found to have a deleterious effect on reasoning, and surprisingly both groups endorsed *fewer* fallacies on emotional statements.

To find no impairment in reasoning with emotional content in the DPD group is in line with the predictions. This suggests that the emotional blunting often reported by people with DPD can have the advantageous effect of preventing the cognitive biases that can often occur with emotional material. There was even a tendency for the DPD group to show superior reasoning on these items. Perhaps there is something about the statements that enables more accurate reasoning, if they are not perceived as emotionally charged. One possibility is that they have more 'real-world' content and are somehow more concrete than the neutral statements, and that this provides an advantage ordinarily obscured by the physiological component of emotional content. Further experimental manipulations are needed to test this idea, as the current paradigm does not explore this issue.

The fact that people with AS performed better on with emotional content is also puzzling. The difference is slightly less marked in this group because they tended to endorse fewer simple fallacies in the first place (although not significantly); yet the difference between their performance on statements with emotional and neutral content remains considerable. One explanation is that the non-global cognitive style preferred by people with AS also allows them to process the statements independently from their emotional content. If people with AS do not recruit semantic memory when processing such statements, then perhaps they also avoid activating 'emotional' memory stores. If this is the case, then the statements will not be tagged as 'emotional', and the biases ordinarily seen will not occur. Again, however, it is impossible to say from these data alone, and further experimental tasks are needed to explore this possibility.

There were certain limitations to this study. The lack of a healthy control group prevented proper statistical analyses, and the reliance on normative data restricted this further. Moreover, the unequal number of statements for some comparisons - i.e. emotion vs. neutral, necessitated the use of percentages, and hence, non-normally distributed data which may have affected the findings. It may be worth replicating this study focusing on emotional vs. neutral statements and additional vs., simple statements, in line with the predictions and hence allowing the use of raw scores.

To conclude, these data reveal that neither clinical group has a global reasoning deficit. On the contrary, both groups excel on particular aspects of these tasks and it seems that the cognitive style displayed by each clinical group can lead to advantages in certain domains. This is of importance in demonstrating intact cognitive function, and suggests that widespread cognitive impairment is unlikely to play a causal role in the interpretation of subsequent data.

3.3 Physical problem-solving task

Autistic spectrum disorders may represent an extreme form of the male brain (Baron-Cohen, 2002a, Baron-Cohen, 1999a, Baron-Cohen and Hammer, 1997). The Extreme Male Brain (EMB) hypothesis sprang from a long tradition of research into gender differences in cognitive processing. The idea is that there are two main ways of understanding the environment: 'empathising' and 'systemising'. Empathising is the drive to identify the thoughts and feeling of other people, and respond emotionally. Systemising, however, is the tendency to analyse the components of a system in order to deduce the rules that govern it; a system being something with an input, a rule-governed operation and an output (Baron-Cohen *et al.*, 2003b).

The controversial element of this hypothesis is the proposition that men are more likely to lean towards 'systemising' and women towards 'empathising'. However, the proposal is that men and women have a *propensity* towards a particular kind of processing, rather than an absolute way of seeing the world. This theory also allows for the existence of a cognitively balanced brain which does not have a preference for either viewpoint. Several studies suggest that the different genders vary in their cognitive approaches (Baron-Cohen, 1999a, Baron-Cohen, 2002a) although the precise cause of this variation is difficult to discern. There is some evidence to suggest that the systemising/empathising distinction is at least in part genetic (Baron-Cohen, 1999a, Baron-Cohen, 2002a, Lutchmaya and Baron-Cohen, 2002), but this is not conclusive.

More specifically, the relationship between the two set of skills, and whether or not they are independent mechanisms, remains unclear. In one sense, they seem to be on a continuum with some kind of trade-off between the two, and the pattern observed in ASDs supports this (see below). A small but significant inverse relationship has been found between the EQ and the 'Systemising Quotient' (SQ) (Baron-Cohen *et al.*, 2003b); the latter offers a self-report measure of systemising - see below. However, people exist that are good (or bad) at both skills, suggesting they are distinct mechanisms (Baron-Cohen *et al.*, 2003b). This issue is important to the EMB hypothesis, and the nature of the interrelation between these two domains requires further empirical investigation.

But how does this relate to people with ASDs? People with ASDs are thought to be at the extreme end of the 'male' continuum and hence very good 'systemisers'. This proposal stems in part from the observation that people with autism do not show deficits in theory building *per se* despite being bad at constructing a 'theory' of mind. There is evidence that their performance in the domain of physics equals, or even surpasses that of people without ASDs. For instance, children with autism can appreciate the workings of a camera (Baron-Cohen *et al.*, 2003a). This observation came from tasks designed a decade ago as a control for 'false belief' tasks (see chapter 1) they match in cognitive load and structure. A child is shown a scene containing an object and is asked to take a photograph of it; while the photo is developing the object is moved and the child asked where the object will be in the photograph. Despite the fact that the photograph is no longer congruent with the visual scene, children with autism tend to get this kind of task correct. The one crucial difference between this task and a false belief task is that a camera is mechanical as opposed to intentional. This finding, coupled with an abundance of anecdotal evidence regarding the obsessions of some people with ASDs towards mechanical systems, led to the formal proposal of an extreme male brain hypothesis of ASDs.

More recently, Baron-Cohen, Wheelwright, Stone, & Rutherford (1999b) tested 3 people with high-functioning autism on a physics task, and on a test of 'empathising' that involves reading another's emotional state merely from a picture of the eye region (Baron-Cohen *et al.*, 2001). All 3 participants were found to be bad at the latter while excelling at the former. The sample were selected on the basis of their superior skills in the fields of maths, physics and computer science; all subjects which fall into the domain of systemising, so their superior performance on the physics test was no surprise. However, it wasn't clear how they would perform on the eyes task, and their actual performance was consistent with the EMB theory and supports the notion of a trade-off between the two domains.

As autism is likely to have at least some genetic origins, a series of studies has also been conducted examining the professions of the parents and grandparents of people with ASDs. A study of over 1000 families of people with ASDs, with a high response rate of 919, found that the fathers and grandfathers are more than twice as likely than 4 control groups (Tourette's Syndrome, Down's Syndrome, Language Delay and Non-language delay) to work in the field of engineering (Baron-Cohen *et al.*, 1997), with 28.4% of the ASD sample having at least one father or grandfather who was an engineer. This study was further replicated with a smaller American sample (n=79) (Mearns, 2002), and a similar pattern of results was found.

Lastly, a recent paper has looked at the relationship between 'empathising' and 'systemising'; as measured by the EQ and the newly designed 'Systemising Quotient (SQ). The SQ has been constructed to complement the EQ, insofar as it has the same number of items, response options, filler items and reverse-score items. However, the content differs, focusing on 'examples from everyday life in which systemising could be used to varying degrees' (Baron-Cohen *et al.*, 2003). Findings confirm male superiority on the SQ and female superiority on the EQ – in line with the EMB theory. In addition, people with AS scored significantly more highly on the SQ than matched controls and lower on the EQ, again consistent with the EMB theory.

In order to examine this theory and to confirm that both of our clinical groups are capable of reasoning outside of the emotional domain, participants were given the 20-item physics test used in the study by Baron-Cohen *et al.,* (1999b).

The predictions were that:

- The AS group will score higher than control participants in the original study.
- The AS group will outperform the DPD group.
- There will be an inverse relationship with scores on the EQ.

Procedure

The physics test was given to participants as part of the wider testing session. For the AS group, it was given in between the two sections of the conditional reasoning tasks detailed above. This was for the purposes of consistency with the DPD group' but also to break up the conditional reasoning task which can demand a lot of attentional resources. The DPD group likewise, received the test in between the two sections of the conditional reasoning task but interspersed with the emotion vignettes, that will be detailed in Chapter 4.

Analysis

The mean scores for each group can be found in Table 3.6 (again these are stratified by gender). All 16 participants from each group filled the questionnaire, although there were 7 instances of missing data (out of 320) resulting from only 3 participants' questionnaires. To prevent losing the 3 datasets in their entirety, the missing data were replaced with the median score for that item, and so resulting in 2 items being scored as correct and 5 as incorrect.

Table 3.6 Mean and SD scores on the physics test for each group

		sex				Group Total	
		male		female			
		Mean	SD	Mean	SD	Mean	SD
group	DPD	11.7	1.8	9.8	2.5	11.2	2.1
	AS	11.2	3.1	9.5	2.1	11	3
Group Total		11.4	2.6	9.7	2.2	11.1	2.6

Independent t-tests were also run to explore gender differences. The groups were collapsed due to lack of variation. There were no significant differences between men and women's scores ($t = 1.55$, df 30, $p > .05$) possibly due to the unequal sample sizes (women $n = 6$).

A one sample t-test using the data from Baron-Cohen *et al.*, (1999), whereby the mean score was 13.2, was performed on the data from men in both of the clinical groups. The analysis was conducted using just the male data to be consistent with the normative data which were collected from 14 men. Significant differences were found between the male DPD participants and the control group (t = -2.9, df 11, p =.014) and between the male AS group members and controls (t = -2.4, df 13, p = .033).

Further analyses were conducted to explore the relationship between the total EQ score and performance on the physics task. For the AS group there was a negative correlation between performance on the physics test and total EQ score (r = -.486, p = .057); again with no significant association within the DPD group (r = -.155, p > .05), although this finding was in the negative direction.

Discussion

Despite the fact that both clinical groups performed significantly worse than controls on the physics task, their scores were not considerably lower. It should be noted that the analysis was restricted to a one sample t-test, hence not utilising the full spread of scores in the original normative data. Furthermore, the participants in the original study had full-scale IQs, as estimated from the NART, of 130.9 ± 6.4 as opposed to 117 (± 5.9 in DPD group and 6.4 in AS group) for the clinical groups in the current study. The authors of the original study recognise the role of IQ in this test, and it seems therefore that high IQ may have confounded their results to some extent. Taking these factors into account, it is reasonable to conclude that both clinical groups in this study can reason effectively in the physical domain. This supports the mounting evidence that neither clinical group suffers from a generalised reasoning deficit.

A non-significant trend was observed for gender differences with male superiority on this task. Unequal sample sizes may have attenuated the significance of this finding which was in the predicted direction. Nevertheless the trend is line with the EMB theory; especially regarding the females in the DPD group, who would be expected to be lower on systemising ability. However, these data alone do not shed any light on the causality of the findings.

The most interesting finding from this study is the moderate inverse relationship between total EQ score and performance on the physics task for the AS group. Although this correlation was only borderline significant ($p = .057$), the magnitude of the association is much stronger than that found between the SQ and EQ (Baron-Cohen *et al.*, 2003b). This may be of help in understanding the nature of the empathising/systemising trade-off, as the data suggest the trade-off to be more prominent with particular aspects of systemising, namely technical or physical systems – at least for people with ASDs. The fact that no such relationship was found in the DPD group does not detract from this finding. Presumably the proposed trade-off between these two skills would not hold for people who have a potential deficit in empathising, which if it occurs is acquired after childhood and fluctuates (Baron-Cohen *et al.*, 2003a), as opposed to being developmental and enduring. Such a deficit would be expected to be independent of 'systemising' skills.

To conclude: full statistical comparison with the control data set would be of value, as this would permit more stringent matching of the clinical and control groups on possible confounding or interacting variables such as IQ, age and gender. Nonetheless, these data go some way to support the Extreme Male Brain theory of autism and more importantly for the current study they confirm that the participants in the clinical groups have yet another cognitive domain intact. At least with the AS group and perhaps with the DPD group, these results coupled with the rest of the data presented in this section, suggest a move away from seeing these conditions as disabilities and towards characterising them as 'differences' or cognitive styles (Baron-Cohen, 2002b).

4. Conclusion

The clinical groups described in this chapter have been diagnosed according to standardised and stringent criteria by qualified psychiatrists. In addition, valid and reliable 'gold standards' have been administered where possible. All in all, these measures confirm that the clinical groups in question are homogenous with regard to their diagnosis and representative of the populations from which they are drawn.

Demographically, both the clinical groups and the control group are drawn from a varied socio-economic background with different ages and a range of occupations. Apart from the AS group, they came from a varied mix of nationalities and ethnicities. In the control group, gender was evenly distributed but there was a lack of women in both the DPD (n = 4) and the AS group (n = 2) which must be considered in subsequent analyses. All groups were well matched for verbal IQ as estimated from the NART all groups having a mean verbal IQ that is above average.

According to self-report measures, the AS group reported significantly lower empathy levels and higher 'personal distress' levels than the control group. However, item-by item analysis revealed that differences on the 'emotional reactivity' subscale of the EQ may be in part due to problems employing a representational theory of mind rather than the other components of 'emotional reactivity'. The DPD group, by contrast, did not score any lower than controls on this particular dimension of empathy, despite predictions. They did, however, show a non-significant trend towards lower levels of 'personal distress' than the other groups. The DPD group also reported problems applying the skills required to show competence in social interactions.

In terms of cognitive functioning, both clinical groups were shown, in more than one domain, to be able to reason at the same level as, and sometimes superior to, controls. The DPD group outperformed healthy controls on conditional reasoning tasks with emotional content. The AS group also showed this pattern along with enhanced performance in tasks that benefited from reasoning in a piecemeal fashion. This was in line with predictions, and may highlight another domain where the cognitive style of people with AS, i.e. weak 'central coherence', is a bonus. Both clinical groups also performed within the normal range for their IQ on the physics task – confirming yet another spared cognitive skill. For people with AS, this was inversely related to their empathy skills and may therefore help untangle the relationship between systemising and empathising.

To conclude, neither clinical group displayed a global reasoning deficit. This means that any data generated by emotional reasoning tasks in subsequent chapters are unlikely to be attributable to basic, low-level deficits. The AS group displayed impairments with the representational aspects of theory of mind in line with previous literature, although the full nature of their deficits in 'affective empathy' remains unclear. The DPD group, on the other hand, did not show a global empathy deficit as measured by self-report, although the findings suggest they may have problems with applied empathy and self-orientated emotional reactions to others' mental states. The next step is to design objective empathy measures to confirm the subjective reports of both of the clinical groups.

Novel and Objective Measures of Empathy

1. Background

Two types of empathy have been distinguished, namely cognitive and affective (see chapters 1 & 2). Although cognitive empathy can refer to all types of mental state attribution i.e. epistemic, desire-based and emotion – affective empathy is most likely to result from observing other people's emotional states. Furthermore, there is a large body of literature on the appreciation of non-affective states but little on the comprehension of affective states, especially from more complex non-body based stimuli (see Chapter 1). To address this issue, a novel task was designed to focus on the attribution and appreciation of affective states from real-life diary extracts. The protocol was designed to capture the two main components of empathy: cognitive and physiological.

Contents at a glance

Specifically, the use of the cognitive representation of the self was measured. This is a factor which may distinguish between the two competing philosophical approaches to mental state attribution: i) 'theory-theory' [the ability to deploy theoretical knowledge based on a set of principles (Zahavi and Parnas, 2003)] and ii) 'simulation' [the capacity to imagine being in the same situation as another (Decety and Chaminade, 2003)]. In addition, emotional arousal in response to the affective states of the protagonist was measured.

1.1 The appreciation of affective states

The attribution of affective states has only really been examined in the context of simple paradigms tapping lower-level, more basic processes (Blakemore and Decety, 2001b). These studies have relied mainly on facial emotional expressions alone, using both the whole or part of the face (Adolphs *et al.*, 2001a, Adolphs *et al.*, 2001c, Dailey *et al.*, 2002, Dolan *et al.*, 2001, Kucharska-Pietura *et al.*, 2003, Baron-Cohen *et al.*, 2001, Hariri *et al.*, 2002b, Lidaka *et al.*, 2003, Gur *et al.*, 2002, Adolphs, 2002, Shaw *et al.*, 2003a, Jansari *et al.*, 2000), and/or voice (Hornak, 1996a, Hornak, 1996b, Morris *et al.*, 1999), including various experimental manipulations such as the mode of presentation (Prkachin, 2003). More challenging tasks have also been constructed based on 'morphed' faces which vary the intensity of the displayed emotion (Surguladze *et al.*, 2003). Recently, Adolphs and Tranel (2003) manipulated the presence and absence of the face in emotion perception and found that participants with selective brain damage could label the emotions in the absence of the face despite not being able to when presented with the face alone. This suggests that the perception of facial affect may be neurally dissociable, in part, from the appreciation of emotional states from other bodily expressions (see chapter 7).

More complex mental state attribution tasks which recruit higher level cognitive processes (Blakemore and Decety, 2001b), invariably focus on epistemic states so that the meta-representational ability of the mind can be tapped unequivocally. As a result, little is known about the processes involved in different types of affective state attribution and even less is known of how they may fractionate. A few studies have focussed on emotional state attribution from more complex stimuli, but mainly in the context of online studies, attempting to discover the brain basis rather than the underlying and accompanying cognitive processes (see chapter 7).

A series of studies have looked at memory processes for emotional events using both images and narratives (Blakemore and Decety, 2001a, Brierley *et al.*, 2004 ~~in-press~~) but the stimuli used in these studies are not specific to other people's emotional states. However, Decety and Chaminade (2003) examined emotion attribution using video-clips of actors telling sad and neutral stories and manipulated the congruency of the motor expression with the content of the narrative. This was a Positron Emission Tomography (PET) study and the resulting neural correlates were shown to be consistent with a 'shared representations' network (see chapter 1 and 7). However, in terms of the behavioural correlates, likeability ratings (which were taken as an index of 'sympathy') were significantly higher in the congruent emotional condition e.g. sad story/sad face and were lower in the incongruent emotion condition e.g. sad story/happy face.

Some data exist on the perception of emotion through the voice (see above), however, studies using written material alone are rare. This is partly because the majority of real-life emotional communication relies to some extent on either non-verbal or vocal processes. However, the content of language is also an important source of emotional cues, and as Decety *et al* (2003) showed, concordance between content and expression is important. To develop tasks that use lexical stimuli may be helpful as a symptom of some clinical conditions may be a specific inability to decode emotion from facial arrays (see below, chapter 1 and chapter 5). However, it studies based on written vignettes do have their own drawbacks, not being completely generalisable to everyday life whereby communication is often face-to-face (Parkinson and Manstead, 1993).

Theoretically it is possible to make some predictions about the kind of processes that may be involved in vignette based tasks. Higher level cognitive processes, similar to those invoked in traditional ToM paradigms, may be recruited (theory-theory and/or simulation - see chapter 1), depending on the level of meta-representation required. It is also possible that 'lower level' processes may be employed to comprehend some aspects of affective states, similar to those involved in the attribution of intention from action (Blakemore and Decety, 2001b). It is also possible that non-specialised processes, such as those based on semantic memory and categorical knowledge, may be used; this may be especially so for stimuli that are very easy to decode i.e. 'children are happy when they go to parties' (Baron-Cohen, 1991). It is important to control for this possibility, as any clinical group with an emotion processing deficit may be adept at employing such skills as the basis of a compensatory strategy. For this reason, care needs to be taken to empirically measure the processes underlying emotion attribution, especially if the tasks are to be of use in furthering the understanding of clinical groups.

In developing such tasks, it is also important to manipulate the difficulty of the task in order to prevent a ceiling effect and prompt people use empathic processes rather than simply rule-based reasoning. In addition, it is vital that the tasks be as ecologically valid as possible in order to try and recreate the empathic response in an experimental setting (Parkinson and Manstead, 1993).

1.2 Physiological arousal

Most emotion attribution studies fail to consider whether physiological arousal occurs when perceiving another person's emotion[16], although there are many studies exploring affective responses to other types of stimuli (Phillips *et al.*, 2003). Emotion induction appears intuitively to be a component of emotion perception. Developmentally, an imitation mechanism with accompanying physiological arousal, would be the obvious candidate for the transfer of knowledge between persons about the qualitative aspects of emotional states (Meltzoff and Gopnik, 1993a). Newborns are known to imitate facial expressions and this is thought to be an early mechanism 'for understanding that others are 'like me' [which] underlies the development of ToM and empathy for others' (Meltzoff and Decety, 2003). One recent pivotal study suggests that imitation is still an important component of emotion perception in adults. Carr *et al.*, (2003) asked participants to simply observe emotional expressions in one condition and imitate them in another, and found similar brain areas were involved with greater activation in the imitation condition. However, it is still unclear whether the imitation was accompanied by emotional arousal.

A further study of relevance here was conducted by Adolphs *et al.*, (2000) found that people with lesions in the right somatosensory cortices were more likely to have problems with emotion perception. The authors interpret this finding as consistent with

[16] Except when this is explicitly built into the experimental design i.e. part of an emotion induction procedure – see Weiss, U., Salloum, J. and Scneider, F. (1999) Correspondance of emotional self-rating with facial expression, *Psychiatry Research*, **86**, 175-184.

'the idea that we recognise another individuals emotional state by internally generating somatosensory representations that simulate how the other individual would feel when displaying a certain emotion expression'. They further point out that such simulation could be overt or covert and there is no need for the individual to consciously feel the state of the other. These findings suggest that emotion perception does involve an element of emotion induction, albeit covertly. However, this still needs to be tested experimentally.

Taylor et al., (2003) had participants rate their own emotional state whilst viewing pictures with affective content as an adjunct to the main experimental manipulation. The authors found that during sad images (which were the only images requiring the participant to empathise) self-reported emotional arousal was reduced when participants had to rate the pictures rather than simply passively view them. This suggests that any physiological arousal may be dampened by task demands. However, the subjective emotional ratings occurred after the task-related ratings and it is possible that this effect would be minimised if arousal was measured before the task commenced. Davis *et al* (1987) also asked participants to rate their own mood after watching dramatic film stimuli and found that emotional reactions of different valences were mediated by different components of empathy.

As mentioned in chapter 1, EMG measures have also been used to explore facial muscle reactions (Cacioppo *et al.*, 1988) and recent studies suggest that these effects are not simply due to demand characteristics but genuinely reflect experienced emotions (Dimberg *et al.*, 2002). Studies have also relied on skin conductance, heart rate measures and the startle response to measure physiological reactions in emotion processing paradigms (Lang *et al.*, 1993, Vrana, 1988, Pissiota *et al.*, 2003).

However, it is rare to find these methods employed in studies using other people's emotional states as the stimuli. Although, Decety and Charminade (2003) in the study described above, measured skin conductance and found a significant interaction response during the condition where facial expression and story content were incongruent for emotion, but no effect during the narrative or motor expression conditions alone. The potential problem with all these measures, and in particular with self-report scales, is that they are explicit and hence may be open to demand characteristics[17] i.e. subjects respond in the way they believe is expected of them. This is especially important with a construct like empathy which is a socially desirable trait (see chapter 3).

One implicit but indirect method of measuring physiological arousal is speech rate. Vocal expression is complex, varied and has numerous components including intensity, loudness, pitch, frequency and tempo (Cook, 2003). There are many studies examining the perception of affective states via prosody (Kucharska-Pietura *et al.*, 2003, Ross *et al.*, 1997) but less using vocal cues as an index of the emotional state of the speaker. Anecdotal evidence suggests that we speak faster when scared or anxious and more slowly when sad or depressed. This assumption finds some empirical support in a study which explores the correlates of expressive behaviour when recounting personal events. Participants showed more cardiovascular arousal when speaking about personal events in a mood-congruent way compared to a mood incongruent speech rate (Siegman and Boyle, 1993). Furthermore, Breitenstein *et al.*, (2001) systematically manipulated speech rate and pitch for 5 emotions and found a faster rate to be associated with fear and anger and a lower rate to be associated with

[17] Although some studies have employed independent raters rather than EMG muscle measures – see Ibid.

sadness. A review of emotion in voice and music confirmed these patterns, and found that articulation rate also increased when the speaker is happy (Scherer, 1995, Cook, 2003). However, it does seem that speech rate is more sensitive to arousal than it is to valence, but this is the case with most physiological measures and speech rate has the distinct advantage of being implicit.

A few studies have attempted to use speech rate as a measure of physiological arousal. Teasdale and Fogarty (1979) tested normal volunteers using mood induction and a counting task to assess speech rate. They found significant differences for mood on speech rate. In attempt to reduce the variability and demand characteristics, McKenna and Lewis (1994) asked participants to count both at normal speech rate and as fast as they could, before and after a mood induction procedure. They found that mood induction affects both maximal and self-paced speech rate, with speech rate accelerating in the happy mood condition and decelerating in the depressed mood condition. Lastly, Barrett and Paus (2002) used a mood induction technique with 49 volunteers and examined several speech parameters, including rate. Demand characteristics were controlled for by using 'neutral' sentences as the reading stimuli, and distracting participants from the true nature of the task. A between-participants design was used to control for carry-over effects. Differences were found between the sad and happy mood condition, with participants having slower, quieter and more monotonous speech when sad.

The utility of the speech rate method is also supported by studies that employ it as an indirect measure of depressive symptomatology. Barrett and Paus (2002) make tentative links between the 'hypopfrontality' observed in depression and the relationship between affect and this particular motor behaviour. Other authors have linked speech differences in depression to dysfunction in meso-limbic nigrostriatal areas (Flint *et al.*, 1993). Teasdale *et al* (1980) conducted a series of case-studies using speech rate and self-report measures of depression and despondency i.e. the Beck Depression Inventory (Beck *et al.*, 1988b). They found that deceleration of speech rate was significantly associated with increased unpleasant affect whether resulting from depressive symptomatology or experimental manipulation i.e. mood induction. Godfrey & Knight (1984) used various indices of speech rate to assess depression levels in clinical groups from admission to discharge and found significant differences over time, suggesting speech rate to be a good outcome measure for treatment for depression.

In the current study we sought to replicate these effects using the affective states of others' as the source of emotion induction. If the participant is in a state of parallel empathy, it is expected that their speech rate will accelerate after the participant has processed the diary extracts containing 'fearful' and 'happy' situations and decelerate after the extract containing 'sad' material. Anxiety levels and 'personal distress' were also measured, as both of these may affect arousal. In addition, participants were asked to perform this task before they were required to explicitly process the stimuli i.e. answer the questions (Taylor *et al.*, 2003). The relationship between this measure and the different components of empathy was considered, particularly in response to vignettes of different valence (Davis *et al.*, 1987).

1.3 Use of the 'self-concept'

The ability to take the other person's perspective is a prerequisite of any empathic response. Traditional 'theory of mind' (ToM) tests rely on the ability to take the perspective of the target, and in so doing, set aside one's own perspective -

particularly with representational epistemic states. There are two major philosophical theories as to the nature of mental state perspective-taking namely: Simulation and Theory-Theory (see above and chapter 1). However, this issue has largely remained in the domain of philosophy and there is a lack of empirical work distinguishing between the two. Social psychologists have also been interested in perspective-taking, and being slightly distanced from the ToM programme, their work is from a slightly different theoretical basis.

Davis (1996) examined the use of the self-concept in perspective-taking. In this task participants watched a videotape of a person speaking about personal events and were asked to 'imagine the self' or 'imagine the target' in the situation or in the control condition, simply watch the video. Participants in the 'imagine self' and 'imagine other' conditions attributed a higher proportion of traits that they had previously attributed to themselves to the character, than those who had instructions to simply watch the video.

There is a growing body of empirical work examining the cognitive and neural correlates of the self and self-processing (Kelley *et al.*, 2002, Kircher *et al.*, 2002, Kircher *et al.*, 2000, Kjaer *et al.*, 2002). However, there is little work examining the use of the self-concept in mental state attribution. Vogeley et al (2001) examined the possible neural correlates of self- and other-processing using an adapted ToM paradigm based on false belief vignettes. The same vignettes and probes were presented in both the third and first person and both common and distinct areas of activation were found. The authors interpreted this as suggesting that neither the 'simulation' nor 'theory-theory' account is likely to be mutually exclusive. Consistent with this interpretation are a further set of studies which have found traditional ToM areas to be activated when participants reported self-generated thoughts (McGuire *et al.*, 1996), when they reported a self-induced tickling sensation (Blakemore *et al.*, 1998) and when reflecting on their own abilities, traits and attitudes (Johnson *et al.*, 2002).

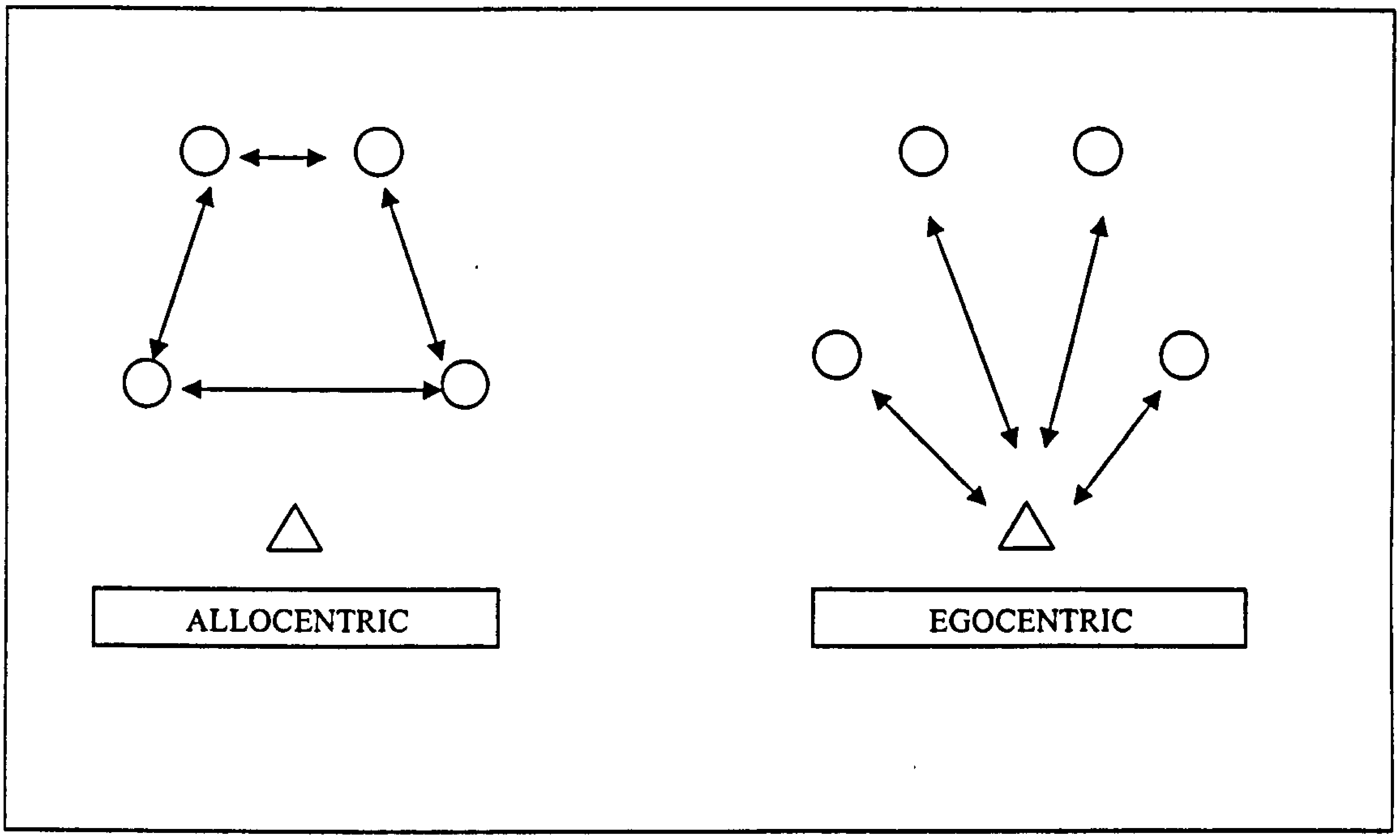
Although Davis does not explicitly affiliate his work with the philosophical theory of 'simulation', there are considerable parallels. It is reasonable to suggest that the main difference between a 'theory-theory' and 'simulation' approach is the use of the 'self-concept' in understanding the other's mental state (Vogeley *et al.*, 2001). In essence, a person using the 'simulation' approach would reason along the lines of "I like parties so John likes parties", the theory-theory approach would, however, predict the use of a more objective strategy such as "people like parties so John likes parties". Although, both approaches are more complex than this, in terms of solid predictions, it seems likely that if a person is 'simulating' the mental state of another, they will use and have available a cognitive representation of the self (see chapter 1 for an overview).

In addition, Davis (1994) also points out that perspective taking can prime people for at least two different types of affective state; i) sympathy and compassion, and/or ii) feelings of unease and personal distress. Much has been written by social psychologists regarding the distinction between these two outcomes, however, this is rarely considered within the neuroscientific or philosophical literature, despite the fact that it is crucial to the development of an empathy model. Not only does care need to be taken to distinguish different types of affective response but there is also the possibility that these distinct responses are based on different cognitive processes. This generates questions open to empirical measurement; does 'personal distress' result from the same kind of cognitive process as 'empathy' or can perspective-taking be further fractionated?

The first step in answering these questions is to empirically establish the relationship between the self-concept and perspective-taking, aside from the resulting emotional responses. There are a few studies that may be of relevance here. Davis *et al.*, (1996) found a linear relationship between perspective-taking and self-concept as measured by trait overlap, so that an increase in perspective-taking led to an increase in overlap between the traits attributed to the target and self (or vice versa). However, he only tested healthy volunteers and one recent study, examining schizotypy within the normal population, suggests that perspective-taking may differ between groups.

Langdon and Coltheart (2001) tested people who scored both low or high on schizotypal traits on both visual and mental perspective-taking tasks. They found that the people scoring high on schizotypal traits not only performed badly on the mental perspective-taking task but also on certain aspects of a visual perspective-taking task. The visual task required the participant to rotate the self relative to an array (as opposed to rotating the array relative to the self). In order to explain these results, the authors borrowed concepts from the field of spatial navigation namely: ego- and allocentric frames of references (see Figure 4.1).

Figure 4.1 Allocentric and egocentric frames of reference adapted from Vogeley & Fink (2003)



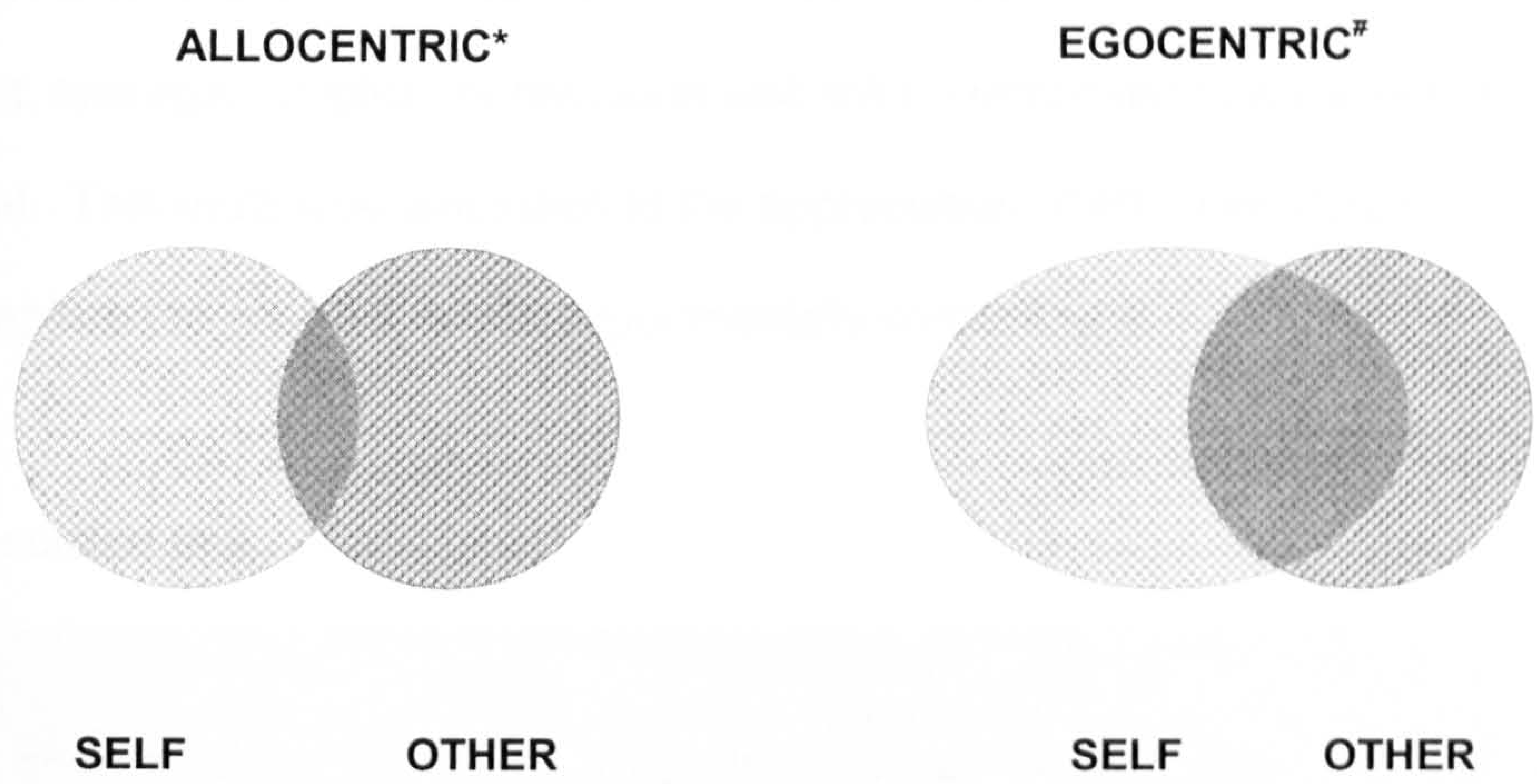
The authors define allocentric simulation as that the ability to reconcile one's own first person view of the world with other people's viewpoints and *on an equal footing*[18]. Egocentric simulation, however, is described as being tied to one's own point of view. Similarly, Vogeley and Fink (2003) define an allocentric viewpoint, as a framework that is independent from the agents position (see diagram 1). However, they also point out that such different mental representation skills may not be best explained with reference to the same concepts as those employed to define abstract body representation. The main finding in this study is that people scoring high on schizotypal traits may excel at egocentric simulation at the cost of allocentric

[18] This is not to be confused with a detached theory-theory approach as it is still simulation and so is likely to involve the use of the self-concept. It may be useful to consider 'allocentric' as 'putting yourself in another's shoes' vs. 'egocentric' as 'putting the other in your shoes'.

simulation and that this underlies their difficulty with certain aspects of visual and mental perspective taking. This effect was replicated in people with schizophrenia (Langdon *et al.*, 2001). There is some evidence that the overlap between self and other is unlikley to be absolute in healthy controls (Ramnani and Miall, 2004) and others also point out that too much overlap could lead to conceptual confusion (Decety and Sommerville, 2003) ~~and~~.

The idea that shared perspectives can be egocentric is not new. Feningstein and Abrams (1993) found that people are more likely to assume that 'others' think the same way them, when there is an increase in self-focussed attention. But this has to be balanced with the known increase in empathy than can occur when the target is in a similar situation to the self [for instance see Houston (1990)]. However, it does seem probable that perspective-taking may further fractionate, and this may be related to the *extent* to which the self-concept is utilised. It seems reasonable to suggest that an egocentric frame of reference would rely more heavily on the self-concept than an allocentric frame of reference (see figure 4.2). This, in conjunction with the existence of distinct emotional responses i.e. personal distress and empathy, may indicate that the association between the self-concept and perspective-taking may not be simply linear. Rather it may tail off once an optimum level is reached – whereby the use of the self-concept beyond that would be indicative of 'egocentric' perspective-taking.

Figure 4.2: Predicted degree of overlap between self and other representations



* Possible basis for empathic emotional response

\# Possible basis for personal distress response.

This idea may have consequences for the pattern of results expected from the trait task (Davis *et al.*, 1996) especially with different groups. For instance, an egocentric frame of reference may be preferred by people who score highly on self-absorption or narcissistic traits and/or in groups who have difficulty adopting the third person perspective. In terms of the current study, it is important to replicate the self-trait overlap effect in a different sample of healthy volunteers, before testing the hypothesis that different groups may vary in their use of the self-concept. It is also important to explore the relationship between empathy, personal distress and other demographic variables with the trait overlap task.

To summarise, the aim of the set of tasks described below is to measure basic affective state attribution including the use of the 'self' concept in this process and the physiological arousal that may accompany it. An implicit cognitive task will be employed to examine the use of the cognitive representation of the 'self' and an indirect, and again implicit, behavioural task will be employed to measure physiological arousal. This multi-level approach to the appreciation of affective states is designed to fully capture the process, whilst experimentally controlling it.

It is predicted that:

- Performance on the diary extracts will be associated with 'empathy' scores on the EQ, with a higher accuracy rate associated with higher empathy rates.
- Each vignette will be accompanied by speech rate indicative of *congruent* physiological arousal.
- The extent of self and other trait overlap will be related to 'empathy' scores on the EQ

2. Task development

A novel task was developed with 3 main components; affective state labelling, speech rate and a trait task. As these tasks are new, their development will be described in depth.

2.1 Vignettes: stimuli design

The task consists of 4 vignettes which were adapted from real diary extracts displayed on a website called www.opendiary.com (see appendix 3). Real diary extracts were used to increase ecological validity (Parkinson and Manstead, 1993). There are 3 vignettes with affective content and one tapping epistemic states. Situations which were considered likely to evoke certain emotional states, were used i.e. birth, death, and common fears such as injections. At no point were the target emotions explicitly stated, instead they have to be inferred from the text as there is evidence that people tend to naturally infer emotions when reading text with emotional content (Gygax *et al.*, 2003). The protagonists were two men and two women and the age range was varied.

The extracts were followed by 3 multiple choice questions requiring participants to identify the emotion displayed by the protagonist. For each question there were 4 alternatives from which to choose. To prevent the test from being too easy, the alternatives were slight variations on the correct answer, often less or more extreme versions of the same or similar affective states or semantic opposites in the control vignette (see below). The three questions were also graded in difficulty; i) basic emotion i.e. sadness in response to the death of a family member ii) basic emotion but less obvious reactions to the situation i.e. anger in response to a death iii) more complex mixed emotions (Plutchik, 1980) i.e. guilt in response to a death.

Gygax et al (2003) found that that experimentally people tend to infer broad rather than specific emotions from written text which may add to the difficulty of this task and prevent ceiling effects. The control vignette included first, second and third order beliefs and participants were also probed as to their own opinion on the subjects covered, in order to tap their ability to represent other peoples' minds. As an additional control, each vignette also had a question designed to tap memory for the facts in the extract.

Format of target questions:

1st Level Belief – what a character thought

1st Level Emotion – basic emotion

2nd Level Belief – what a character thought another character thought

2nd Level Emotion – basic emotion stemming from 1st level emotion

3rd Level Belief – what an character thought a second character thought about the first character.

3rd Level Emotion – primary mixed emotion (Plutchik, 1980).

Procedure

The vignettes were presented in written format, rather than orally, to control for vocal expression. Participants were given the vignette to read and then they were instructed to turn over and answer the questions without referring back. Standardised instructions were given:

> Below you will find some questions with a choice of 4 answers. Please read the question and say <u>out loud</u> the answer which you think is most accurate – there is no need to write it down. <u>Please do not refer back to the extract</u> and work as quickly as possible as you will be timed. You do not need to read the whole answer out – you can just say the letter – For Example; 'No 1 – A, No 2 – B etc'.
>
> Remember it is important that you answer all the questions – sometimes more than one answer might seem appropriate – if this is the case then please give the answer that fits the best. Or if you think none of the answers are correct then please give your best guess.

Responses were timed to the 0.01 second using a stopwatch from the moment participants gave their first answer until their last answer.

Piloting

In the initial pilot, 10 people (7 women, 4 men, mean age 30.6 yrs, ± 4.1) completed 4 emotion vignettes and 4 belief vignettes, which were counterbalanced. On the basis of these data the most appropriate belief extract was selected for further use. The text in the vignettes was also amended where necessary, as was the word length. The second pilot involving 5 people (3 women, 2 men, mean age 31.4 yrs ± 7.4) resulted in some further minor amendments in the text and word length of the extracts.

Final Vignettes

All the vignettes were further controlled for length and ease of reading level using Microsoft Office XP Professional – Word 2002, and the statistics for the final vignettes were:

Emotion A

Flesch Reading Level 8.78	Questions 7.92
Total Word Count 327	Vignette 173 words & questions 139 words

Emotion B

Flesch Reading Level 8.00	Questions 7.35
Total Word Count 316	Vignette 169 words & questions 143 words

Emotion C

Flesch Reading Level 5.5	Questions 4.84
Total Word Count 307	Vignette 170 words & questions 137 words

Belief

Flesch Reading Level 7.25	Questions 6.89
Total Word Count 308	Vignette 167 words & questions 141 words

To control for the practice and other order effects, the final vignettes were counterbalanced using a 4 x 4 Latin Square. With additional non-arousing tasks (traditional ToM vignettes – see Chapter 6) between the diary extracts to further prevent carry over effects. The additional tasks were also counterbalanced using a reversed 4 x 4 Latin Square (see chapter 6). To prevent a response set, whether or not the correct answer was the most extreme of the alternatives was also varied.

2.2 Counting task

The counting task was used as a measure of speech rate and as an indirect measure of physiological emotional arousal. Participants were instructed to count out loud from 1 to10 after each of the vignettes. As it was important that this test remain an implicit measure of arousal, demand characteristics were carefully considered. To this end, the counting task was also given after some different vignette-based tasks (traditional ToM tests – see above) to prevent participants realising that it related solely to the diary extracts. In order to obtain baseline counting rates, the counting task was also given once before all the vignettes and again after the diary extract without emotional content.

Piloting

See above for demographic details. The main effect was that people count more quickly as they move though the task. As the vignettes were counterbalanced, fatigue and order effects were already controlled, so this did not pose any additional problems.

2.3 Trait task

The paradigm used by Davis *et al.,* (1996) was adapted for use in the current study. Rather than relying on freely generated trait lists, the 'Trait Adjective Checklist' (Loehlin and Nichols, 1976) as used by Davis *et al.,* (1996), consisting of 159 traits presented in alphabetical order, was used in the pilot study.

Piloting

Participants were sent the list prior to the experimental session and were given it again after each vignette. The instructions were:

Self

Which of the following adjectives do you consider to be descriptive of yourself? Place a tick beside any adjective that you might use in describing yourself to someone else. Your behaviour will vary with the situation, of course, so place a tick beside the adjectives which might apply to you frequently, even though they are not appropriate all the time. Work rapidly putting down all your first thoughts.

Target

Please place a tick next to the following adjectives that best describe James. We appreciate that you only have limited information but we are interested in your initial thoughts and there are no wrong or right answers. So please work rapidly putting down what first comes to mind.

The main change made from the pilot study was a reduction in the number of traits used. The list was pared down to 75 adjectives by using the traits that were endorsed most frequently whilst ensuring an even balance of positively and negatively-valenced traits (see appendix 4).

3. Method

The final version of the full task was validated on a group of psychologically healthy volunteers.

3.1 Participants

Volunteers were recruited via a circular email at the Institute of Psychiatry (IOP), a database of volunteers maintained by staff at the Institute and postcards in local newsagents, local library and job centre. There were 48 participants who were a subset of the 53 reported in chapter 3 (the data from the first 5 of the original sample were included in a pilot run - see above). Of the 48 participants, 26 (54%) were recruited from academic or clinical staff at the IOP and the remaining 22 (46%) people were recruited from the local area and non mental health workers at the IOP. The occupation status of the sample was therefore varied.

The sample consisted of 23 (48%) men and 25 (52%) women. The mean age was 34.38 years (± 11.23). Forty-two (88%) reported their nationality as British and 1 (2%) Sri Lankan, 2 (4%) American, 1 (2%) Irish, 1 (2%) German and 1 (2%) Indian. Thirty-six (75%) reported their ethnicity as white European, with 12% indicating that their ethnicity differed from their nationality being 2 (4%) afro-carribean, 1 (2%) black, and 2 (4%) mixed race.

Six (13%) of the group were educated to secondary school standard, 4 (8%) to further education, 13 (28%) to higher and 24 (50%) to postgraduate level. 21 (44%) people were single, 10 or 21% co-habiting and 10 (21%) married and 3 (6%) people divorced with the remaining people either separated, widowed or with a non-cohabiting partner.

3.2 Design

This study is quasi-experimental as EQ score, which is based on a naturally occurring trait, is the independent variable. A repeated measures approach was used for all the variables.

3.3 Procedure

Volunteers had the details of the study explained to them before deciding whether or not to participate. Before the main testing session, the EQ, a social desirability scale (Crowne and Marlowe, 1960) and stage 1 of the trait task were sent to participants along with some additional measures not reported here.

The volunteers attended in person for between 35 minutes and an 1 hr for which they were paid £5.00. They were given details of the study and a chance to ask questions before signing the consent form. The tasks were administered in a standardised way and the whole session was tape recorded, ostensibly for accuracy, but also as a means of testing one of the dependant variables: speech rate.

3.4 Measures

3.4.1 Diary extracts

In the absence of confirmation from the protagonist, mental state attribution is to some extent arbitrary. The diary extracts (see above and appendix 3) were therefore scored as 1 for the 'preferred' answer and 0 for 'other' answers, based on the pilot data responses. However, one of the questions consistently evoked a different answer from 'high empathisers' (ascertained from a median split on the EQ) than that given in the pilot data. This answer was therefore coded as the 'preferred' answer, contrary to the pilot data. As there were 3 diary extracts depicting affective states each with 3 multiple choice questions – the total score was out of 9. The questions on the extract portraying belief's were coded as '1' for a correct answer and '0' for an incorrect answer, and so the total score was out of 3.

3.4.2 Speech rate

Speech rate was measured using a desk-mounted microphone and a portable tape-recorder. Participants were asked to count from 1 to 10, without an explanation as to why and post-experimental questioning revealed that many of them thought it was a distracter task and no-one guessed its real purpose. This therefore overcomes previous criticism of demand characteristics. The speech rate was then measured directly from the tape using a stopwatch which timed to 100th second.

Coding – as speech rate is known to vary between individuals this variable was calculated within-participants. The time taken to count from 1-10 was converted to a percentage of the baseline speech rate for each participant. The decision was made to use the control vignette speech rate as a baseline, as this was considered a more natural index of speech rate than the first instance when people were asked to count.

3.4.3 Trait task

Participants were presented with a list of traits and asked to tick those that best described themselves. The list was given at least 24 hours before the experimental session although in most cases the interval was much longer. During the experimental session, participants were then asked to tick from the trait list, the traits that best describe the character in the extract.

Coding – the number of traits that the participant had previously endorsed for themselves and later attributed to the target was calculated and a mean was then calculated for all 3 affective state vignettes. According to Davis *et al* (1996), there are two ways of calculating the proportion of overlap: i) percentage of the 'self' traits that were later endorsed for the target (self-overlap) and ii) percentage of a 'target' traits that were earlier endorsed for the self (target-overlap).

4. Results

All datasets were first examined for normality and where deviations were found the datasets were transformed using the appropriate method. Simple analyses were then performed for on the data gathered from each task, to explore the association between the outcome variables (accuracy rate on extracts, percentage trait overlap, speech rates) and demographic factors (see section 3.1): age, sex, verbal IQ as measured by the National Adult Reading Test (NART), occupation - mental health worker or not, education which was dummy coded, and self-report measures; total score on IRI personal distress items, and EQ factor scores (see chapter 2).

Variables that had a significant linear relationship with the outcome variables were entered into a multiple regression analysis to control for potential confounds that may be problematic in a quasi-experimental design. So for each dataset the following analyses were conducted where appropriate:

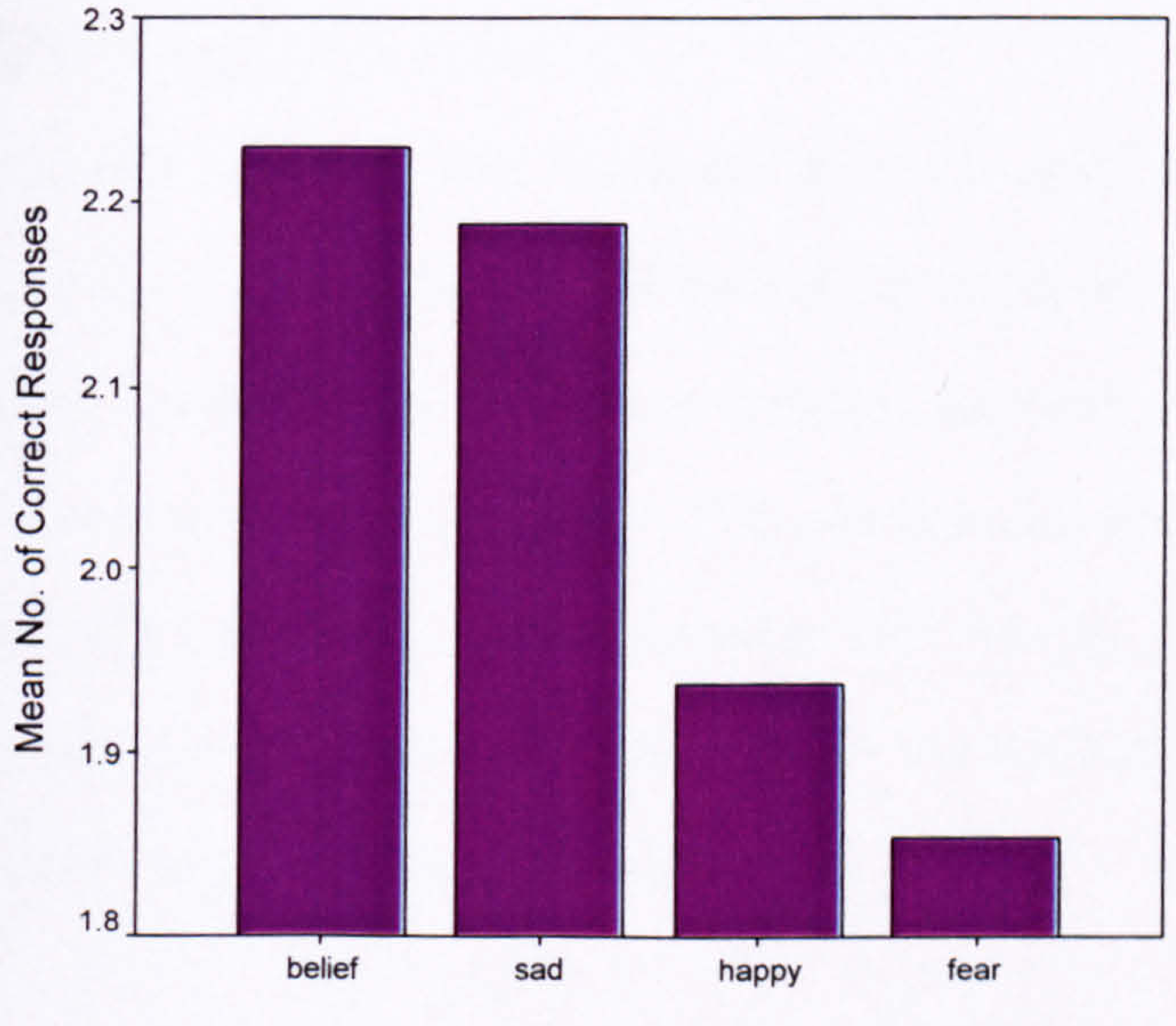
- Simple analyses examining the association between the outcome variable, demographic factors and self-report measures.
- Where significant associations were found, multiple regression analyses were conducted with the variables of interest.

4.1 Affective state labelling

Descriptive analysis

The mean number of correct responses out of 3 for the control vignette was 2.23 (± .69), the 'sad' vignette was 2.19 (± .79), the 'happy' was 1.94 (± .86) and for the 'fear' vignette was 1.85 (± 8.5) – see figure 4.3.

Figure 4.3: Mean number of correct responses on the mental state labelling task

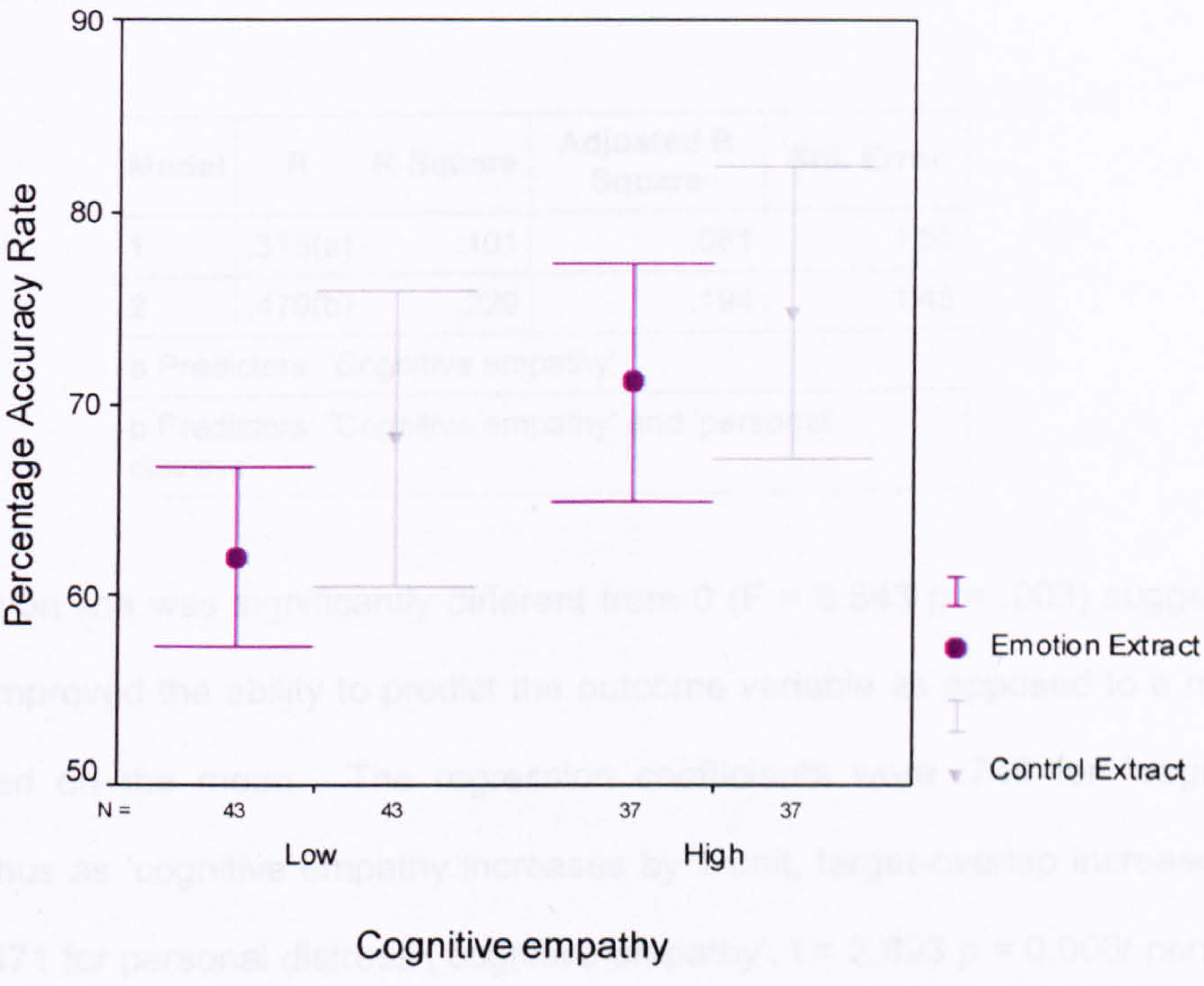


Demographic factors and self-report measures

A Pearsons Product Moment Correlation showed positive relationships between accuracy rates on the diary extracts, factor 1 'cognitive empathy' (n = 48, r = .313, p = .030) (see figure 1) and total score on IRI items (n = 47, r = .288, p = .047). - tests also revealed differences between those who had been educated to postgraduate level and those who had not (t = 1.194, df 46, p =.058).

People high on 'cognitive empathy' also performed more accurately on the *control* vignette (see figure 4.4), although this relationship was not significant. It is unsurprising that 'cognitive empathy' which includes items relating to all classes of mental state would also be associated with epistemic mental state attribution. However, the pattern observed is also consistent with the fact that more 'affective state' than 'epistemic state' items loaded onto 'cognitive empathy' (see chapter 2). No significant associations were found between any other demographic variables or EQ factors and the control vignette with non-affective content.

Figure 4.4: Mean percentage accuracy rate on diary extracts with 95% confidence intervals



Multiple regression analyses

Cognitive empathy' and 'personal distress' were both significant predictors of accuracy on the diary extracts containing affective states (see Table 4.1). There was a moderate and positive correlation between both these variables and the outcome variable (Multiple R = .479) which accounted for 22.9 % of the variance in accuracy rate in this sample (adjusted R^2 = 19.4%). This suggests that these two variables explain approximately 1/5th of the variation on this task and that other factors must also play a substantial role.

Table 4.1: Summary of predictors of accuracy on affective state labelling task

Model	R	R Square	Adjusted R Square	Std. Error
1	.318(a)	.101	.081	1.55
2	.479(b)	.229	.194	1.45
a Predictors: 'Cognitive empathy'				
b Predictors: 'Cognitive empathy' and 'personal distress'				

The regression line was significantly different from 0 (F = 6.543 p = .003) suggesting the model improved the ability to predict the outcome variable as opposed to a model simply based on the mean. The regression coefficients were .749 for "cognitive empathy' (thus as 'cognitive empathy increases by 1 unit, target-overlap increases by 7.49) and .471 for personal distress ('cognitive empathy', t = 2.893 p = 0.006/ personal distress t = 2.702, p = .010) showing that although each predictor is making a significant addition to the model, cognitive empathy plays the larger role (see table 4.2). This is confirmed by the standardised coefficients which indicate "cognitive empathy" to be the best predictor followed closely by 'personal distress'.

Table 4.2 Coefficients for predictors of accuracy on affective state labelling task

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	Cognitive empathy	.611	.271	.318	2.252	.029
2	Cognitive empathy	.749	.259	.391	2.893	.006
	Personal distress	.471	.174	.365	2.702	.010

Both predictors had a positive association with the outcome variable. There were no significant inter-correlations between the predictor variables satisfying the multicollinearity assumption. These data suggest that both 'cognitive empathy' and 'personal distress' play an important role in labelling affective states.

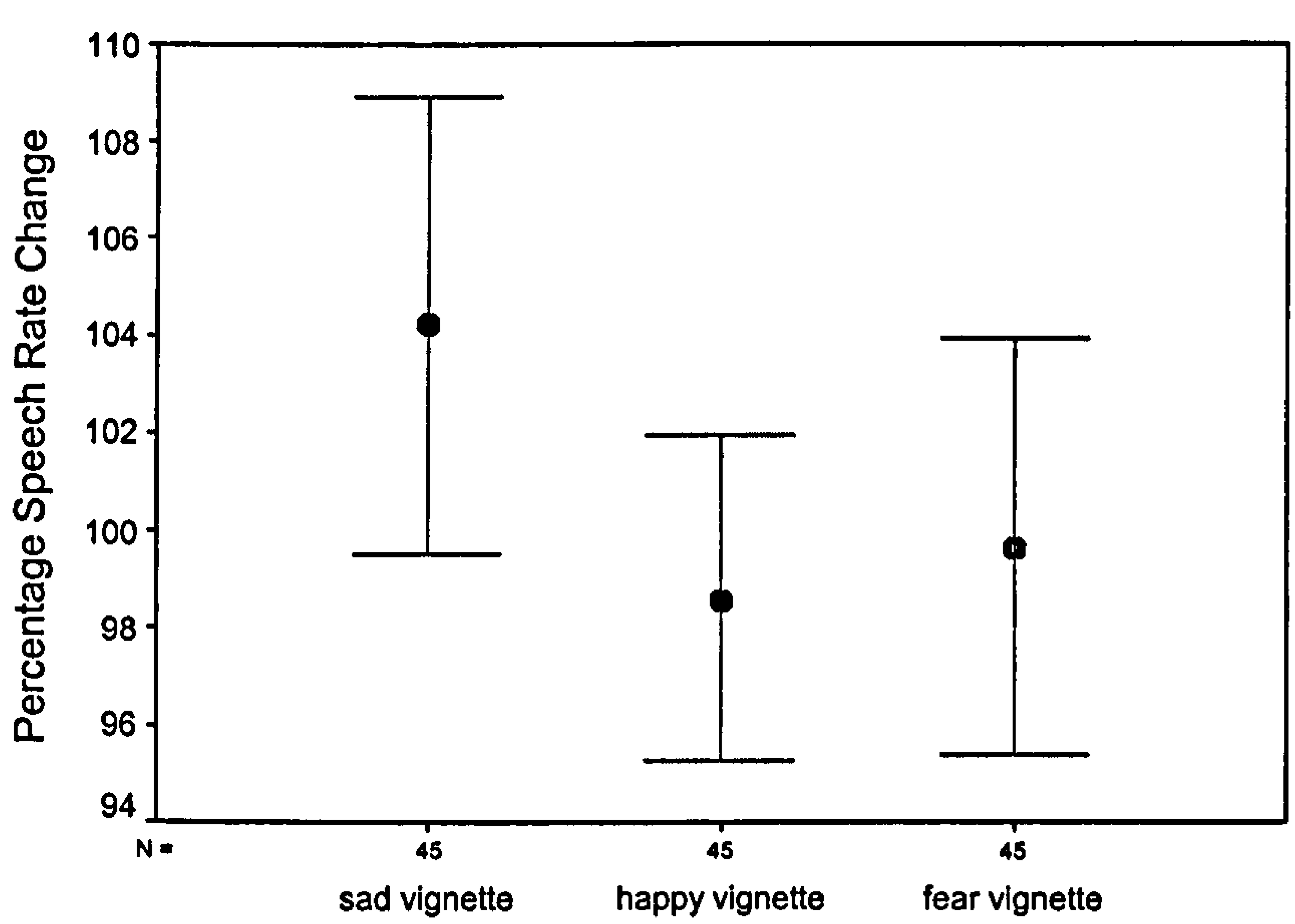
In a separate analysis, none of these variables were significant predictors of performance on the control extract which may suggest different processes underlie the attribution of non-affective states. However, as the control task had a smaller range it is not possible to assert this conclusively.

4.2 Speech rate

The correlation matrix for speech rate and EQ factors showed no strong associations between empathy and speech rate. On the basis of these data, the decision was made to systematically examine the mean speech rates for each vignette, as a percentage of the baseline (see Figure 4.5). The data were first examined for outliers and 2 cases were dropped and the data log transformed.

Paired t-tests revealed a significant difference between the speech rate following the 'sad' and 'happy' vignettes (t = 2.325, df 44, p =.025), with the difference between the 'sad' and 'fear' vignette approaching significance (t = - 1.942, df 44, p =.059) but no difference between the 'happy' and 'fear' vignettes (t = .370, df 44, p > .05) (see Figure 2 which is based on original scale for ease of reference). Hence, as predicted, speech rate decelerated with the mood-congruent vignettes and accelerated with the fearful and happy congruent vignettes, as predicted.

Figure 4.5 Mean speech rate as a percentage of the baseline with 95% confidence intervals



a. Over 100% Represents Slower than Baseline

b. Under 100% Represents Faster than Baseline

Pre- and post-test anxiety levels did not correlate significantly with any of the speech rate variables and neither did scores on the personal distress items of the IRI.

4.3 Traits

Demographic factors and self-report measures

The 'self-overlap' variable was transformed using square root. The relationship between both trait overlap variables: i) 'self-overlap' ii) 'target-overlap' (see above for details of calculations), demographic details and self-report measures, was then explored. There were also 2 additional variables of interest: i) 'mean number of traits attributed to the target' and, ii) 'total number of traits attributed to the self' (see Davis *et al.*, 1996).

Both 'self-overlap' and 'target-overlap' showed a significant association with 'emotional reactivity' (tendency to respond emotionally to others' affective states – see chapter 2) suggesting increased use of the self-concept is related to the ability to respond affectively to others' emotional states ('target-overlap' $n = 47$, $r = .446$, $p = .002$/'self-overlap' $r = .342$, $p = .019$). The 'number of traits attributed to the self' was also related ('target-overlap' $r = .865$, $p < .001$/'self-overlap' $r = .540$, $p < .01$) as was the 'mean number of traits attributed to the target' ('target-overlap' $r = .536$, $p < .001$/'self-overlap' $r = .881$, $p < .001$) which is expected seeing as the 'overlap' variables are calculated on the basis of these figures.

Multiple regression analyses

A multiple regression analysis was performed to examine the contribution of the different variables in predicting the variation in the outcome variable. The number of self traits attributed in the diary extract depicting beliefs did not have any significant associations with any of these variables and so was not analysed further.

With 'target-overlap' (percentage of traits attributed to the target that were earlier endorsed for the self) as the outcome variable in the regression analysis, 'no of traits attributed to self' and 'emotional reactivity' were significant predictors (see table 4.3). The variables were strongly correlated with the dependent variable (Multiple R = .892) and together 'emotional reactivity' and 'no. of traits attributed to self' accounted for 79.5% of the variance (adjusted R^2 .786). The additional variance accounted for by adding 'emotional reactivity' to the model was approximately 5% as the 'no of traits attributed to self' alone accounted for 74.8%.

Table 4.3: Summary of predictors of 'target-overlap'

Model	R	R Square	Adjusted R Square	Std. Error
1	.865(a)	.748	.743	13.5567
2	.892(b)	.795	.786	12.3691
a Predictors: 'no of traits attributed to self'				
b Predictors: 'no of traits attributed to self' and 'emotional reactivity'				

As a whole the model had predictive value over and above that derived simply from the mean ($F_{(2, 45)}$ = 86.35, $p < .001$). The regression coefficient for 'emotional reactivity' was 6.37 (thus as 'emotional reactivity' increases by 1 unit, target-overlap increases by 6.37) and for 'number of traits attributed to self' it was 1.47 and the standardised coefficients revealed that the latter was the best predictor at .783 (indicating that an increase of 1 SD in no of traits attributed to the self results in an increase of .783 of a SD in 'target-overlap). 'Emotional reactivity' was the secondary predictor (.252) – see table 4.4.

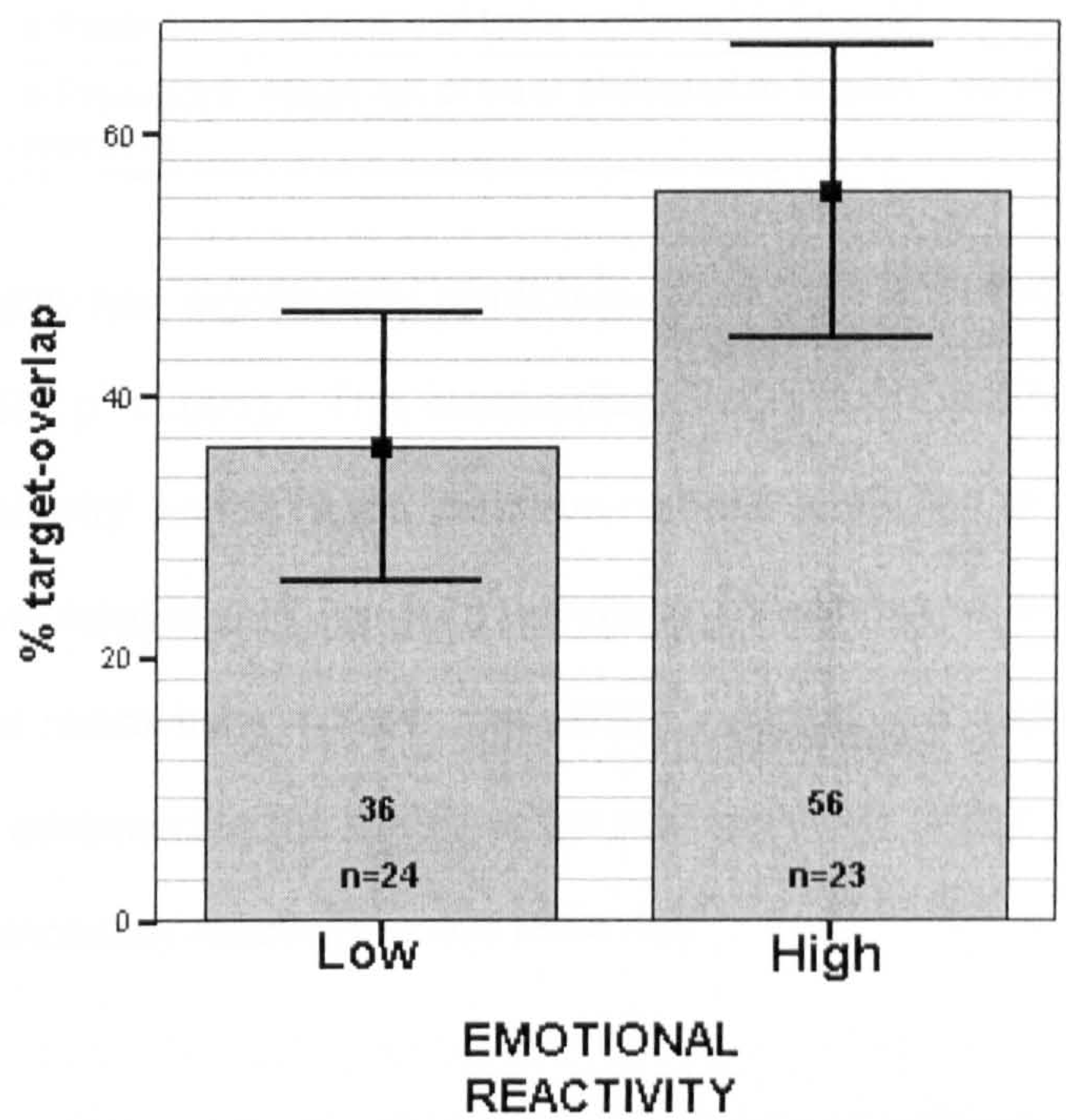
Table 4.4 Coefficients for predictors of 'target-overlap'

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	no of traits attributed to self	1.628	.141	.865	11.569	.001
2	no of traits attributed to self	1.512	.134	.803	11.322	.001
	'emotional reactivity'	5.648	1.781	.225	3.171	.003

This is consistent with the fact that the 'no traits attributed to the self' overlaps with the outcome variable. There was a positive relationship between 'target-overlap' and both 'emotional reactivity' ($t = 3.51$ $p = 0.004$) and 'number of traits attributed to the self' ($t = 10.93$, $p < 0.001$) suggesting that an increase in both 'emotional reactivity' and 'no of trait attributed to self' is accompanied by an increase in trait overlap.

A median split was also calculated on emotional reactivity scores in order to further represent the effect of this variable on trait overlap (see Figure 4.6).

Figure 4.6 Mean percentage target-overlap for low and high empathy groups with 95% confidence interval error bars



With 'self-overlap' (percentage of the 'self' traits that were later endorsed for the target) entered as the outcome variable, 'mean no of traits attributed to target' and 'emotional reactivity' were the only significant predictors. Again, these were strongly correlated with the outcome variable (Multiple R = .904) accounting for 81.8 % of the variance (adjusted R .809). 'Mean no of traits attributed to the target' alone accounted for 77.4 % of the variance again indicating that 'emotional reactivity' explained an additional 5% of the variation in the criterion variable (see table 4.5).

Table 4.5: Summary of predictors of 'self-overlap'

Model	R	R Square	Adjusted R Square	Std. Error
1	.880(a)	.774	.768	.7544
2	.904(b)	.818	.809	.6843
a Predictors: 'mean no. of traits attributed to targets'				
b Predictors: 'mean no. of traits attributed to targets', 'emotional reactivity				

The model again had significantly predictive value than is based on the mean alone ($F_{(2, 45)}$ = 96.59, p < .001). The standardised regression coefficients were .219 for 'emotional reactivity' and .819 for 'mean no of traits attributed to the target' and there was a positive relationship for both variables (mean no of traits: t = 12.10, p < .001/emotional reactivity: t = 3.24, p = .002). Again, this suggests that the 'mean number traits attributed to the target' is the best predictor of the 'self-overlap' variable followed by 'emotional reactivity' – see table 4.6.

Table 4.6: Coefficients for predictors of 'self-overlap'

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	'mean no. of traits attributed to target'	7.543E-02	.006	.880	12.262	.001
2	'mean no. of traits attributed to target'	7.026E-02	.006	.819	12.104	.001
	'emotional reactivity'	.326	.101	.219	3.236	.002

A correlation matrix indicated no significant relationships between 'self-overlap' or 'target-overlap', calculated on the basis of traits attributed in the 'belief' vignette, and EQ or 'personal distress' measures.

5. Discussion

As predicted, the attribution of affective states, as measured by performance on the diary extracts, was related to 'cognitive empathy' according to scores on the EQ. 'Personal distress' as measured by the IRI was also significantly associated with successful affective state attribution. In addition, each extract elicited a speech rate suggestive of physiological arousal congruent to the affective state expressed in the vignette. The extracts with affective content were also accompanied by a degree of trait overlap which was partially related to 'emotional reactivity'.

More specifically, 'cognitive empathy' and 'personal distress' combined explained approximately 1/5th of the variance in performance on the affective state labelling aspect of the task. For both measures there was a positive relationship with correct labelling of emotional states, suggesting an increase on either was accompanied by a higher accuracy rate. As expected, the most significant predictor was "cognitive empathy" followed by 'personal distress' and these two items were independent. Interestingly, the control task which tested the understanding of epistemic states was not predicted by either of these variable; it is possible that this is due to the smaller range of scores on the control vignette. Alternatively, it may be because 'cognitive empathy' had many high loadings for items that tap the appreciation of affective states as opposed to epistemic states. This would be consistent with a dissociation between the cognitive mechanisms for processing of epistemic as opposed to affective states.

These results also suggest that those who had a tendency to feel 'personal distress' in response to other's affective states were able to attribute emotions more accurately. This indicates that although their understanding was self-orientated, they were still 'in tune' with the protagonist. Taken together, these two variables provide a sensitive profile of one aspect of an 'empathic person' i.e. someone who scores highly on 'cognitive empathy' but endorses few 'personal distress' items on the IRI scale, hence, distinguishing between those who react in a self- rather than and other-orientated way.

The main limitation of this task is that it is not possible to conclude definitively, from these data alone, that it taps 'empathy' as opposed to cognitively effortful rule-based reasoning, from these data alone. However, the fact that an empathy measure (the EQ) predicts performance on this task goes some way to support its construct validity and suggests it, at least in part, taps empathic processes. Although we cannot be sure that the cognitive empathy EQ items themselves, do not partially tap rule-based type reasoning too. In addition, 80 % of the variance remains unaccounted for by the model suggesting there are other processes at work. A further limitation is that the labelling of the affective states of our vignettes was to some extent arbitrary. It is difficult to overcome this in empirical research without, for example, obtaining self-reports from the target immediately after a mental state is displayed (Ickes, 1997). Lastly, it may be useful in replication to design more vignettes to provide a wider range of scores (especially with the control vignette) and a more sensitive measure of affective state labelling.

Although there were no associations between EQ or 'personal distress' scores and speech rate, the data suggested that the task was successful in tapping physiological arousal and distinguishing valence. The group as a whole decelerated after the 'sad' vignette, despite the fact that there was a general tendency to accelerate as people moved through the vignettes (controlled for by counterbalancing). Furthermore, participants accelerated after the 'happy' and 'fear' vignettes, again in line with the predictions. This variation between stimuli of differing valence was confirmed by paired t-tests which showed significant differences between the 'sad' vignette compared to the 'happy' and 'fearful' vignette but not between the latter two. The fact that anxiety measures were not significantly related to speech rate further suggests that the changes observed resulted from the experimental manipulation. This is the first known study to find that people react 'physiologically' when reading about other people's affective experiences. Furthermore, this study confirms the utility of the speech rate as an indirect and implicit measure of such arousal.

In criticism, however, this measure of physiological arousal only captures the response associated with 'parallel' empathy and does not allow for the distinction between different types of empathy, for example, 'reactive' responses such as sympathy or compassion. In addition, the lack of association observed between EQ scores and speech rate may cast doubt on the ability of speech rate measures to be sensitive enough to detect individual variation in physiological responses. However, it is also possible that this variation is not as pronounced in the normal population as it may be in clinical groups. In replication, it may be worthwhile using more sensitive 'speech print' measures, as have been employed in other studies in an attempt to detect any individual differences related to personality traits such as empathy. It would also be worthwhile controlling for diurnal variation of mood by testing people at the same time of day, especially as empathy may have temporal variation as a function of mood (Nezlek *et al.*, 2001).

The results of the trait task suggest that the self-concept is indeed recruited during empathy. Both types of trait overlap index were significantly predicted by 'emotional reactivity', and 'mean no of traits attributed to targets' and/or 'no of self traits' – depending on the method used to calculate the overlap. In both instances, this combination of variables accounted for much of the variance. However, this is probably due to the fact that both 'number of traits attributed to self' and 'mean number of traits attributed to the target' were not independent from the outcome variable. This association is expected therefore simply because the more self traits endorsed, or more traits attributed to targets, the greater the probability of some of the self-like traits being attributed to target. However, what is of importance is that 'emotional reactivity' was also a significant predictor of both trait overlap variables, above and beyond the sheer number of either set of traits. This therefore suggests that the extent to which the cognitive representation of the self is available after the attribution of affective states is related to the degree of 'emotional reactivity' (the tendency to react emotionally to others' affective states) experienced. This association supports the concurrent validity for the trait task.

Interestingly, scores on 'personal distress' items did not correlate with either outcome variable, which may indirectly suggest that for this group, the degree of trait overlap is related to allocentric rather than egocentric processes. In further work it will be worthwhile ascertaining the degree of trait overlap that is indicative of egocentric vs. allocentric and where (and if) the use of the self-concept becomes non-optimal. The lack of association between empathy, 'personal distress' and trait overlap generated from the belief extract indicates that the use of the self-concept (and so possibly simulation) may be specific to affective rather than epistemic states.

The main limitation of this task is the use of freely generated traits as opposed to a fixed choice protocol. The latter would eliminate a possible source of response bias in those participants inclined to endorse lots of traits regardless of the task demands. Further measures to differentiate 'egocentric' and 'allocentric' processes, such as manipulating the degree of self-focus prior to an emotion attribution paradigm, would also be useful. Despite these limitations, however, the task appears to provide reliable incidental measures of the cognitive processes underlying the attribution of affective states; and to our knowledge this is the first time this has been attempted.

The suggestion that participants used their self-concept and displayed behaviour consistent with physiological arousal provides further evidence that participants employed *empathic* processes rather than *effortful* reasoning strategies to label the affective states of the protagonist. This adds further weight to the task's construct validity. However, although it appears rule-based reasoning was not relied upon in healthy volunteers, it may be possible to perform well on this task using effortful strategies. Care therefore needs to be taken to use more sensitive indices of underlying processes i.e. reaction times, when testing clinical groups. The assumption would be that empathic responses should be quicker that those based on effortful 'cold cognition'.

6. Conclusions

The data reported here from healthy volunteers suggests that the 3 novel and objective tasks described here, validly and reliably tap the various components of empathy, namely: i) mental state attribution, ii) a simulation process and iii) physiological arousal. These tasks should be useful in exploring the subjective and anecdotal reports of a lack of empathy in AS and DPD (see chapter 3 for details of groups). However, as our sample was not randomly drawn from the normal population regarding education, care needs to be taken to control for this when comparison between groups is contemplated.

There is also the suggestion that people with ASDs may process faces (Teunisse and de Gelder, 2003) and facial affect in a different way to control groups (Hobson *et al.*, 1988). However, the findings are mixed (see chapter 6 for full discussion) and it is even less clear how people with AS fit in here.

Several tasks have measured affective state labelling in ASDs using non-facial stimuli. Most of which also require meta-representation (see chapter 1 for explanation). Castelli *et al.,* (2002) used the Heider & Simmel paradigm (Heider and Simmel, 1944), in which participants view geometric shapes moving on a screen and are asked to infer their intentions. The AS/HFA group gave significantly fewer mentalising answers and showed differential neural activation to controls, but it is difficult to draw conclusions about the group's proficiency with affective states from these tasks alone, as a whole range of mental states were applicable i.e. coaxing, chasing, fighting. People with AS/HFA were also found to perform less well than controls on a task requiring the appreciation of a *faux pas* from a written narrative (Baron-Cohen *et al.*, 1999a) but again it is difficult to tease out their performance on emotional states alone, as a *faux pas* requires a joint appreciation of both the belief and emotional states of others.

Royeurs (2001) found that a group of high-functioning people with autism performed the same as controls on static tests of mental state inference but were impaired on a dynamic naturalistic empathy task. This task required participants to watch a video of 2 actors interacting and every time the tape was stopped (72 times), participants were required to write down their inference as to the mental state of the protagonist. Inferences were coded as correct if they matched or were similar to the actual state reported by the protagonist, making this one of the few studies to use real rather than inferred mental states as a benchmark. However, in another paper on the perception of affect in music, a group of children with ASDs were found to perform at the same level as controls (Heaton *et al.*, 1999). Furthermore, in an article reporting the psychological language employed by children with autism, they were found to be comparable to clinical controls in their talk about affective states (Tager-Flusberg, 1992) but were impaired in social emotions.

One of the first papers to test emotion attribution from narratives, both with and without the ability to appreciate representational states, used a test of the cause of an emotion (Baron-Cohen, 1991). Seventeen autistic, 16 clinical controls and 19 clinically normal children were tested on this task that tapped emotion stemming from situations, desires and beliefs. In the 'situation' condition, the children were shown a doll called Jane and told that Jane is going to a birthday party – and asked how does she feel? In the 'belief' condition the child had to predict the emotion that followed a mistaken belief, i.e. Jane was expecting a particular kind of breakfast cereal, that she liked but was given a different which she didn't like. There were significant group differences on the belief task, as predicted by a deficit in meta-representation, but the autistic children were spared in their knowledge of emotions from situations.

It is unclear whether such data provides real evidence that the children appreciated the protagonist's affective states or whether their successful performance relied on the application of a rigid rule-based strategy i.e. 'children like birthday parties'. For instance, it is possible for a child to simply learn that 'happy' applies to children going to parties with little understanding of the concept of happiness, including aspects such as the subjective dimensions (Hobson, 1991). The authors of another study which attempted to teach children with autism about mental states, concluded that it was likely that the group with ASDs used rules to pass tests rather than a genuine understanding (Hadwin, 1996).

Serra *et al.*, (2002) also attempted to separate out awareness of the cause of an emotional state in a study with a group of children with lesser variants of autism. Storybooks were used to assess various aspects of the children's ToM: emotion recognition, the distinction between physical and mental entities, prediction of behaviour and emotions on the basis of desires and beliefs. They found the children to be impaired on the prediction of emotion from different cues and interpreted this as in keeping with an emotion processing deficit. Controls for meta-representational ability were not reported: a potentially important factor, especially as two of the emotions, on which the children scored particularly badly, were curiosity and surprise; both of which are likely to be based on representational states (Baron-Cohen, 1991). The children performed at much the same level as the control group on the 'happy' vignette but not on the 'sad' vignette. Therefore, it is not clear whether these findings represent a real emotion processing deficit and, as with the previous study, whether or not the test can be solved using rule-based strategies.

In a study utilising both verbal and non verbal cues, Loveland *et al.*, (1997) found that people with autism and pervasive developmental disorder who were not high-functioning, had trouble inferring how a person felt based on what they had said, if the emotion was not explicitly stated. However, this was not the case for the high-functioning groups who tended to base their replies more on non-verbal information. Alternatively, a study looking at emotional memory in a group of people with high-functioning autism (HFA) showed them not to display the same enhancement in the recall of affective material as a control group, perhaps indicating a lack of the expected differential processing (Beversdorf *et al.*, 1998). Overall, there is some debate as to whether observed emotion processing deficits can in fact be reduced to domain general skills such as verbal memory and IQ. Buitelaar (1997) found these two skills to be the best predictors of both emotion recognition (based on faces and narratives) and traditional ToM tasks. It is not clear whether such findings are indicative of the role of IQ ordinarily in such tasks or whether people with both autism and high IQs manage to use their intelligence to 'hack out' alternative routes to solve such problems (Happe, 1994a and see chapter 6 for overview).

Lastly, Dennis *et al.,* (2000) found that a group of high-functioning children with autism could label simple emotions but again had specific and significant problems with social emotions such as those based on deception. The emerging picture seems to be that people with ASDs may be able to understand basic emotions, if the need for meta-representation is minimal (which is not the case for social emotions) although it is not clear whether they use the same processes as controls. It is therefore reasonable to predict that people with Asperger's Syndrome may be able to infer the emotional state of the protagonist from a narrative (when not explicitly stated), but it is also entirely possible that they will use different processes to healthy volunteers i.e. the use of rule-base compensatory strategies. A narrative task such as that described in chapter 4, that excludes meta-representational ability as much as is possible, but is difficult enough to prevent ceiling effects and sensitive enough to pick up on the underlying strategies employed, would appear to be useful in exploring this issue further.

1.1.2 Physiological arousal

Uta Frith points out that, anecdotally, there is reason to believe that people with ASDs have problems regarding only the meta-representational aspect of empathy (Frith, 2003). To use her terminology, she argues that there is no reason to believe that people with ASDs do not experience *instinctive sympathy* i.e. an emotional reaction to other's emotional states. She further explains that any apparent empathy problems may be due to an inability to appreciate the mental state of the other in the first place i.e. when requiring meta-representation, and secondly, in knowing how to react to another's emotional state. She gives several anecdotes of people with ASDs who were moved enough to act as a response to other people's distress. Despite such accounts, there is only a small amount of empirical data to back up these claims, although the studies are fairly consistent.

Blair (1999) in one of the first studies to examine this issue directly, measured the galvanic skin response of a group of autistic children, to both non-person based and person-based emotional stimuli, taken from the International Affective Picture System (IAPS). He found that the ASD group had the same responses as the control group, to both sets of stimuli, suggesting that they could respond emotionally to other people's affective states. However, he also points out that from these data alone, it is not possible to conclude that their emotional response was *qualitatively* the same as in the control group. Galvanic skin response is a crude measure of arousal and valence, since emotional states also include a cognitive component, so it is not possible to say from these data alone whether the people with ASDs had an emotional state that was the same or similar to the control group. For instance, considering the evidence of high personal distress in this group (see chapter 3), further probing would be necessary to ascertain whether this response is likely to be self- rather than other- orientated.

A further series of studies have shown intact emotional responses to non-facial stimuli in ASDs. Shalom, *et al.*, (2003) found similar galvanic skin responses in a group of high-functioning children with autism, to that of controls, as a result of viewing 24 non-facial pictures from the IAPS. These lack of group differences suggest both groups showed similar physiological arousal to the stimuli. There were, however, some differences on the behavioural aspect of the task with the autism group rating the unpleasant pictures as less interesting and the neutral pictures as more interesting, than the control group ratings. The authors interpret the results in terms of an impairment in the system that mediates emotions (amygdala) and feelings (medial prefrontal areas). However, they point out that future replications need to control more adequately for response perseveration and verbal IQ. A further study found no differences in emotional modulation of the startle reflex in response to non-facial emotional pictures, in a group of children with ASDs as compared to controls (Salmond *et al.*, 2003).

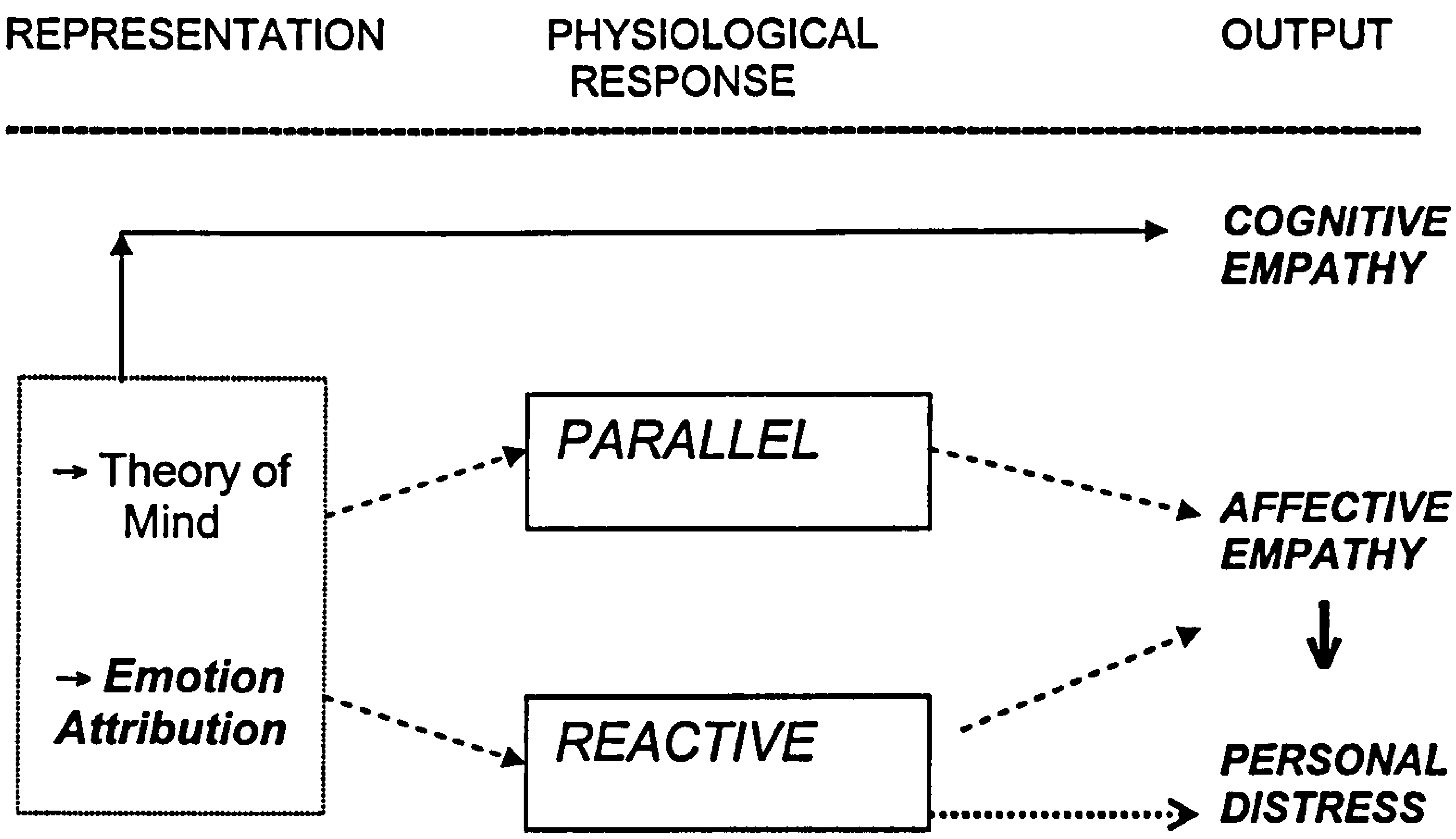
In addition, Yirmiya, *et al.,* (1992) tested autistic children on a task purporting to measure 'affective empathy'. In this study the children were asked to identify a character's affective state from a video and then self-report their own emotional state. Autistic children performed significantly worse than the control group but they still did 'surprisingly well' scoring 6.94/10 as opposed to 8.21/10 in the control group. This suggests that the autistic children were able to respond with parallel 'affective empathy' although not as well as the control groups. Unfortunately, the reliance on self-report alone of one's current emotional state makes this paradigm less suitable for use with adults.

Another set of studies examined the response of children with ASDs to the affective states of the experimenter. Sigman (1992) found a group of autistic children to be less attentive to an adult showing distress, discomfort and fear, than control groups. Furthermore, Kasari (1993) in a study looking at the response of children with ASDs to someone displaying pride towards them, found they showed more avoidance responses, which the authors interpret as indicative of the children finding emotional displays aversive. However, in a study designed to further probe this hypothesis, physiological measures (heart rate) suggested the autistic children did not dislike affective displays (Corona *et al.*, 1998). Instead it seems they paid less attention to the experimenter feigning distress than the control group, despite appearing concerned and interested.

Bacon *et al.,* (1998) examined this issue from a slightly different angle exploring the different responses of those with high and low functioning autism, to an adult simulating distress. He found those with HFA to *respond* in the same way as the control groups i.e. they tended to look at the distressed adult, but they differed from the control groups in their ability to act in a pro-social way i.e. comforting the adult. Those who were low functioning, however, significantly differed on all indices of responding, in line with previous studies. So it seems that at the very least, people with less severe forms of autism may be able to comprehend the distress of other people, although they may not pay as much attention or offer comfort or display other forms of pro-social behaviour, in the same way as controls. Furthermore, it is not clear whether they find displays of affect in others' aversive, as the results are mixed. The 'personal distress' scores from chapter 3 suggest they might.

Regarding self-report measures of empathic emotional responding, it can be recalled that in chapter 3 the 'emotional reactivity' EQ scores of this group were only 1 SD below the control group, much less than the difference on 'cognitive empathy' and 'social skills'. Furthermore, although this difference was statistically significant, an item-by-item analysis revealed this may have been due to items that tap meta-representation rather than empathic emotional reactions. It would be helpful to know, therefore, whether group differences exist on other measures of emotional empathy, with less of a meta-representational load, such as the empathic concern items on the Interpersonal Reactivity Index (Davis, 1980). The fact that this group also scored significantly higher on the 'personal distress' items (see chapter 3) further suggests their emotional responses may be self-orientated (see above and Figure 5.1).

Figure 5.1: Empathy and Related Processes



- Bold italics denote the possible route(s) for people with AS. It is not clear whether parallel affective empathy can be a prelude to personal distress.
- In this instance, theory of mind is used in the restricted sense to refer to meta-representational ability.

It seems therefore, that people with AS are likely to experience affective responses, both to others' emotional states and emotional stimuli in general. However, tests showing this needs to be replicated and further effort needs to be made to establish the nature of the affective response i.e. egocentric vs. empathic. The speech rate task (see chapter 4) provides an implicit and objective means of assessing people's emotional responses. There is evidence to suggest that people with AS find emotion perception from prosody difficult (Rutherford *et al.*, 2002), however, there is no reason to believe this affects their own speech rate. In fact a study of speech and prosody in adults with HFA/AS, found no differences between this group and a comparison group on speech rate (Shriberg *et al.*, 2001). One way of controlling for this and allowing for individual differences would be to calculate speech-rate within-participants and analysis data based on speech rate *change*.

1.1.3 The use of the self-concept

When shown a picture of a ~~smiling~~ frowning face, Christopher Boone said, "I knew that it meant 'sad', which is what I felt when I found the dead dog" (Haddon, 2003). Christopher Boone is a fictional character with Asperger's Syndrome, however, the author's astute characterisation highlights a real possibility that a basic 'online' strategy is open to those with AS. If emotion attribution is in part based on simulation-like processes then this raises the question of whether people with AS could potentially utilise this skill. Whether simulation is a key process in the appreciation of affective states is open to debate, but several recent theories suggest this may be possible (see chapter 1). Furthermore, the control data in chapter 4 suggest that participants used the self-concept in appreciating others' affective states and this is indicative of a simulation-type approach.

As yet, there is no direct evidence as to whether people with AS can simulate mental states. However, there is a body of work on imitation in ASDs and imitation is 'overt action simulation' (Blakemore and Decety, 2001b). Imitation also involves an element of perspective-taking, and may be a precursor to both empathy and ToM (Meltzoff and Decety, 2003). The findings regarding imitation in autism are mixed. Charman and Baron-Cohen (1994) tested a group of 20 people with autism on standardised procedural and gestural imitation tasks and found no differences when compared to a clinical control group. Furthermore, Gepner *et al.*, (2001) observed that some autistic children spontaneously mimicked the emotional expression of an actor in a dynamic emotion perception task. However, several studies have found imitation impairments and the general consensus is that some kind of deficit exists. For instance, Stone *et al.*, (1997) found a group of autistic children to display weaker imitation skills and although they showed a similar pattern of performance across tasks, the imitation of body movements was more difficult than imitation with objects. This independence between different types of imitation is further supported by the results of Hobson and Lee (1999) who observed that people with autism reported imitating the actions of others more easier than imitating the person i.e. their style of action. Recent studies also seem to suggest that people with autism may have 'lower level' problems perceiving biological motion, which may have a bearing on their ability to imitate a person's body movements (Toichi *et al.*, 2002, Blake *et al.*, 2003).

Other findings also suggest an imitation deficit in ASDs (Meltzoff and Gopnik, 1993a, Rogers *et al.*, 2003) including a study with a group of young children (i.e. 20 months) who displayed less imitation then comparison groups (Charman *et al.*, 1997). A recent paper by Williams et al. (2001) has gone so far to suggest that there may be a neurological deficit in the mirror neurone system (see chapter 1) and cite the well-documented symptom of echolalia in autism as evidence. Furthermore, Charman (2002) suggests that an imitation deficit may be the basis of autism as an empathy disorder.

To our knowledge there are no studies specifically examining imitation in Asperger's Syndrome. It is possible that those with AS share the imitation deficits seen in other ASDs. However, there are some differences in the symptomology between AS and ASDs which may be of importance. Specifically, the significantly lower prevalence of echolalia in AS (Szatmari *et al.*, 1989) and the sparing of language (DSM-IV, 1994) especially as tentative links have recently been made between brain areas that subserve imitation and language (see chapter 7 and Hamzei *et al.*, 2003, Heiser *et al.*, 2003 ~~(in press~~)). Furthermore, one recent study with 4 people with AS and one with HFA, using magnetoencephalography (MEG), found activation in the primary motor cortex whilst viewing another's actions to be comparable with controls (Avikainen *et al.*, 1999). This may mean that imitation skills are spared in AS. Avikainen *et al.*, (2003) explored imitation in 6 people with AS and 2 people with high-functioning autism and found no advantage for imitating people as in a 'mirror-image', in the clinical groups as compared to controls. The authors suggest this may be due to a dysfunctional 'mirror neurone' system. However, in the 'non-mirror' image or 'crossed' condition i.e. imitating left side of the body with left side of one's own body etc, there as no difference between groups, which could alternatively suggest some preservation within this system.

It is already known that people with ASDs have difficulties in perspective-taking when the target mental states are representational, although the situation is unclear regarding emotional states. However, there is one more class of perspective-taking that may be of interest, in the visual domain. Children with autism perform as well as controls on such tasks as predicted by a meta-representational account, as such perception relies on primary representation and it is possible to use mental rotation strategies to solve such tasks (Baron-Cohen, 1993). If people with AS apply different strategies to take another person's visual perspective albeit from a largely egocentric stance, then they may be able to mentally simulate from an egocentric perspective too[19], if the target mental state is transparent (see chapter 1). The difficulty still lies in measuring empirically whether or not a person has fully appreciated the affective state including all its dimensions, rather than simply using a rule-based strategy.

The high 'personal distress' scores of people with AS in response to other's distress (see chapter 3) suggest a more comprehensive understanding of others'

[19] It must be remembered that this is not entirely egocentric as it is based on simulation which requires a concept of 'self' and 'other' – see chapter 4.

affective states, than a purely semantic one, but from an egocentric perspective. Furthermore, there is a small and growing literature focussing on self-concept in ASDs. Toichi (2002) found that 18 high-functioning people with autism did not benefit from the self-reference effect in memory in comparison to a control group. Participants were given a list of personality traits and asked either phonological, self-referent or semantic questions about them, and this was later followed by an unexpected recognition test. For the autistic group, self-referent processing did not lead to the expected enhanced memory performance. Furthermore, Gallup *et al.*, (2003) in discussing the link between self-recognition and social cognition, point out that there is some evidence to suggest people with autism 'develop self-recognition later in life and [self-recognition] may be absent in 20 – 30% of cases'. In addition, it has long been noted that children with ASDs can have problems employing personal pronouns (for example see Lee and Hobson, 1994).

Frith and Happé (1999) and others, argue that the concept of self is in fact reliant on ToM and that there is a direct relationship between the two. The nature of self-awareness in ASDs is the subject of debate and many of the issues relating to awareness of other people's mental states are of relevance, including the dissociation between affective states and those requiring meta-representation (for overview of philosophical debate see Zahavi and Parnas, 2003). For instance, Baron-Cohen (1989) suggests that although people with ASDs are 'unaware of their own mental states', this may not include affective states. Frith and Happé (1999) examined self-awareness in AS empirically (n=3) and found that participants indeed had difficulty relaying information about their own mental states and that this was related to their performance on ToM tests. If this is the case then it would be reasonable to predict that people with AS would have a less developed self-concept than controls (although more so that those with other ASDs as their ToM ability tends to be more developed). As Frith and Happé (1999) state, it is as if 'the self isn't central to their world – it is their world'. Frith (2003) further suggests that these problems with self-concept may be at the root of the core deficits in autism.

Consistent with this idea, one study found a group of people with AS to score significantly higher on a self-consciousness scale designed to measure 'attention to the private aspects of the self, such as feelings or motives' (Blackshaw *et al.*, 2001). Frith (2003) further suggests that when a full ToM first develops it is applied to the self, which might explain why the autobiographical narratives of some people with AS are incredibly detailed. It is also possible that this additional insight may also lead to increased vulnerability to social anxiety which high-functioning people with autism may suffer from (see Gillott *et al.*, 2001 for discussion, Amaral and Corbett, 2003).

Adolphs (2001b) found that people with ASDs were able to fully understand at least 88 personality traits and attributes, presented lexically, in the same way as controls. The self-other overlap trait task described in chapter 4 is therefore ideally placed for testing these claims. Based on previous literature, the prediction is that if people with AS are capable of simulating the affective states of others then it is likely to be from an egocentric standpoint. This would also fit with their inability to imitate in a mirror-image fashion despite being able to do so in a crossed way (Avikainen *et al.*, 2003). It seems therefore that people with ASDs may use either an egocentric simulation-like process in order to appreciate the affective states of the protagonist or a rigid rule-based strategy. If a simulation strategy is employed then this will be represented by significantly greater overlap between self and other traits than that observed in control groups.

1.2 Depersonalisation Disorder

People with depersonalisation disorder (DPD) often report a subjective lack of empathy (see chapter 1, 2 and 3). However, just as there are few data based on complex emotion attribution in healthy individuals, data on the performance of those with DPD on such tasks is non-existent. Furthermore, the empirical literature on DPD as a whole is sparce, with comprehensive studies documenting its prevalence as a primary clinical entity only appearing in recent years (Lambert *et al.*, 2001a, Sierra and Berrios, 2001, Simeon *et al.*, 1997). Despite this, it is possible to make some predictions on the basis of recent work and anecdotal reports.

1.2.1. Affective State Labelling

There is no reason to think that people with DPD will be particularly bad at labelling mental states. As a group they are not known for their social incompetence. They are also fairly insightful as to their own symptomatology and do not report any problems in this domain. Furthermore, no labelling abnormalities have been noted in any of the experimental studies. Phillips *et al.*, (2001) and Sierra *et al.*, (2002) found that people with DPD were able to correctly label pictures from the IAPS according to their valence while their subjective levels of arousal were lower than controls. Moreover, their performance on a labelling task (see chapter 6) and scores on cognitive empathy ~~in~~ (chapter 1) were intact. If one accepts that phenomenologically, people with DPD share some of the characteristics as those with psychopathy[20], specifically a lack of some subjective emotional responses, then the latter group are known to perform *well* on ToM type tasks requiring the labelling of mental states (Blair, 1996).

[20] There are important differences between the two conditions, for instance, DPD is acquired in later life whereas psychopathy is a developmental disorder, and this has an impact on nature of the condition. Furthermore, people with DPD complain of a lack of empathy unlike people with psychopathy.

1.2.2 Physiological arousal

The emotional blunting reported by people with DPD leads to the prediction that their physiological responses to other people's mental states will be dampened. However, these anecdotal reports need to confirmed objectively. One of the first studies to look at emotional blunting in DPD found differential neural activation to aversive scenes in contrast to both a healthy and clinical control group made up of people with obsessive compulsive disorder (Phillips *et al.*, 2001). The DPD group had reduced activation in the emotion-sensitive areas of the brain such as the insula. However, this study alone does not provide objective evidence of emotional blunting as there may be all sorts of reasons for differential neural activation in clinical groups. A further study measured the galvanic skin responses (GSR) responses in 15 people with DPD, 15 healthy volunteers and 11 people with anxiety disorders (Sierra *et al.*, 2002). They were shown 5 neutral, 5 pleasant and 5 aversive pictures. Again, statistically significant differential responses indicative of emotional blunting were found in the DPD group when processing the aversive pictures. Furthermore, lowered noradrenaline levels consistent with autonomic blunting were found in a pilot study of 9 people with DPD (Simeon *et al.*, 1998).

Turning now to the self-report questionnaires detailed in chapter 3, a complex pattern emerges. The DPD group's mean scores on the 'personal distress' items were in line with affective dampening. Strikingly, however, the DPD group did not report themselves as impaired on 'emotional reactivity' as one would predict. However, this may have been because these items also tap aspects of cognitive empathy on which this group is not thought to be impaired. It seems therefore that in order to investigate this further, implicit and objective tests of physiological arousal, specifically in response to others' emotional states, are needed. The speech rate measure detailed in chapter 4 was therefore be employed.

1.2.3 Use of the self-concept

Again, there is little empirical work on the nature of the self-concept in DPD, despite the diagnosis and definition of the condition resting on disturbances in the concept of the self. Depersonalisation disorder is defined in the DSM-IV (American Psychiatric Association, 1994) as an 'alteration in the perception or experience of the self so that one feels detached from and as if one is an outside observer of ones mental processes or body' and depersonalisation, in common usage, refers to a subjective sense of unreality in one's self (Hunter *et al.*, 2004 ~~in press~~).

Paradoxically, despite this sense of loss or deterioration of the self, it has been suggested that 'excessive self-observation' may play a role in DPD (Mellor, 1988). Recent cognitive-behavioural accounts of DPD expand on this and suggest that DPD sufferers are constantly preoccupied with comparing how they *actually* feel with how they *should* feel (Hunter *et al.*, 2004 ~~in press~~). The authors also point out that this group has a tendency to high levels of introspection and to over-intellectualise. The key idea is that this excessive focus on the self leads to 'attentional biases that heighten the detection and severity of symptoms'. In support of this argument, the authors also found promising initial results from a treatment program that includes a strategy to reduce self-focus (Hunter *et al.*, 2004 ~~in press~~).

Further support for the role of excessive focusing on the self in DPD symptomology can also be found in the normal psychological literature (Silvia and Abele, 2002). In this study with healthy volunteers the authors manipulated the effect of focusing on the self on the intensity of emotional experience. In the first condition, increased self-awareness was implicitly induced by the use of mirrors and self-narrative, and it was found to decrease the reported intensity of emotional experiences (during mood induction to create both sad and happy moods). When the instructions were to explicitly 'try and experience the emotional state' a reversed result was found, with increased self-awareness increasing intensity of emotional experiences. This highlights the need for implicit tasks that control for demand characteristics.

The evidence of excessive self focus and the anecdotal reports of a lack of empathy, suggest that people with DPD may use egocentric techniques for apprehending other people's affective states. The trait task will be used to examine this and it is predicted that the DPD group will show more self-other overlap than the control group. In addition, a self-report measure purporting to tap self-consciousness in both the public and private domains will be administered (Mckenzie and Hoyle, 2003).

It is predicted that:

- There will be no differences between the control group and those with DPD or AS in the recognition of affective states.
- If the AS group use a different, less 'intuitive' strategy than control group to label affective states, this will be reflected in slower reaction times.
- The DPD group will show a different speech rate pattern than the other two groups, indicative of less congruent physiological arousal.
- Both clinical groups will show significantly more trait-overlap than the control group, between themselves and the character in the diary extract.
- There will be significant group differences on full 'personal distress' scores of the IRI with the AS group scoring higher and the DPD group scoring lower, than the other groups.

- Both clinical groups will positively endorse more items indicative of self-absorption, than controls on the Self-Absorption Scale (Mckenzie and Hoyle).

1.3 Power analyses

Group sizes were based on detecting a conservative effect size (0.2) with 95% power. The effect size was derived from the data gathered in chapter 4. A sample size of 16 for each clinical group (with n = 48 in the control group), with unequal sample sizes, was required.

2. Methods

2.1 Participants

See chapter 3 for details of clinical groups and chapter 4 for details of the control group. Regarding those who returned the full version of the IRI (see chapter 2 for details of controls – p11), there were 10 men and 2 women from the AS group with a mean age of 37.7 yrs (± 10.1), and in the DPD group there were 5 men and 2 women with a mean age of 31 yrs (± 4.7). The same people also returned the SAS scale with 1 additional person from the DPD group returning this alone.

2.2 Measures

The mental state attribution task described in chapter 4 was used. This consists of a series of written diary extracts which convey various affective states without explicitly stating them. Participants were required to choose from 4 options the mental state that best describes the protagonist. This was followed by the speech rate task, again as described in chapter 4. Lastly, participants were required to endorse traits that best described the character in the vignette using the same paradigm as described in chapter 4. In all cases the 'self-traits' were filled in at least 24 hours before or after the experimental session although in most cases the interval was much longer than that.

The full version of the IRI (Davis, 1980) was also sent to participants as was the Self-Absorption Scale (SAS) which is a newly validated questionnaire designed to tap both private and public self-absorption (Davis, 1980). The scale has 18 items, and factor analysis revealed that 9 items load onto private self-absorption and 9 onto public self-absorption. For instance:

- I have difficulty focusing on what others are talking about because I wonder what they're thinking of me – Public Self-Absorption
- My mind never focuses on things other than myself for very long – Private Self-Absorption

2.3 Procedure

All measures were given in a standardised way as reported in chapter 4. The only difference was that people with DPD were not given the ToM tasks and so these were not dispersed in between the diary extracts. Instead, they performed the conditional reasoning tasks (as detailed in chapter 3) between each extract to control for carry-over effects.

As in chapter 4, these tasks were given as part of a wider testing session although some additional tests were administered with the clinical groups (see chapter 3 for details of reasoning tasks which were not administered to the control group) whilst other tests were dropped. Lastly, the full version of the IRI and SAS was sent to all groups a few months after the initial experimental sessions, hence lowering the return rate.

3. Results

Firstly, the dataset for each task was tested for normality and where deviations were found, the data were transformed using either logarithmic or square root. Next, initial simple analyses were performed to compare group performance on each task. This was followed by an exploration of the effect of demographic factors to control for using unmatched groups (see chapter 3). For each task gender, education* (measured categorically), age, occupation (whether or not the participant was a mental health worker) and verbal IQ were considered. Verbal IQ was chosen over other forms of IQ as the tasks used have a major verbal component and because the NART, a verbal task itself, was used to estimate IQ (see chapter 3). Lastly, a full model was performed to examine group differences on the dependant variable for each task, adjusting for any demographic factors previously shown to be of relevance. To summarise, for each task the following analyses were conducted:

- A simple analysis to examine group differences on the dependant variable.
- A simple analysis to explore the association between demographic factors and the dependant variable.
- A full factorial analysis to examine group differences on the dependant variable whilst controlling for the necessary demographic factors.

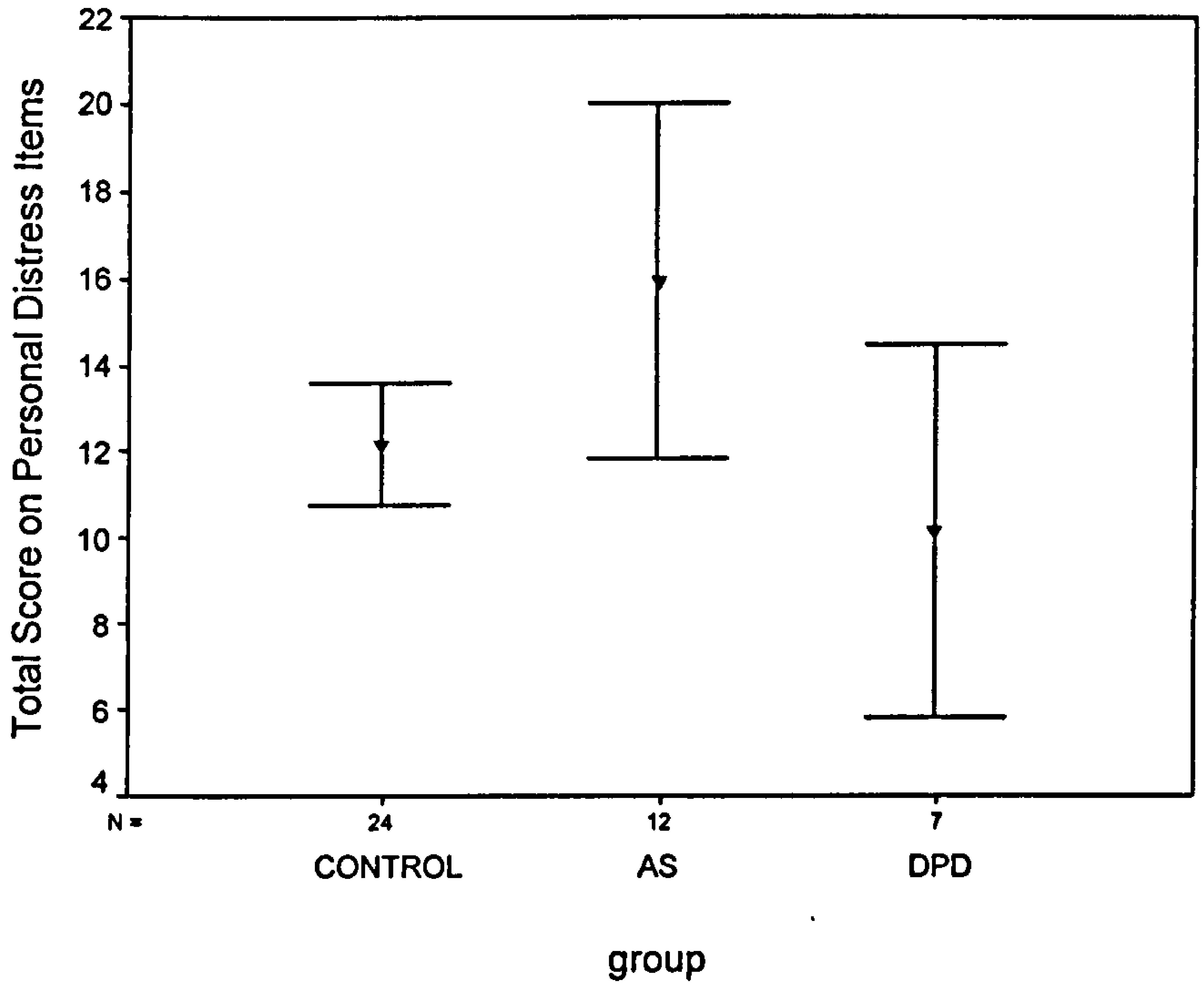
3.1 Interpersonal Reactivity Index and Self-Absorption Scale

Simple group analyses

Full details of the control group scores on the Interpersonal Reactivity Index (IRI - Davis, 1980) and Self-Absorption Scale (SAS - Davis, 1980) can be found in chapter 2. An initial one-way ANOVA comparing all 3 groups on each factor of the IRI showed significant group differences on 'personal distress' ($F_{(2, 42)} = 4.13$, $p = .023$) and 'perspective taking' ($F_{(2, 42)} = 3.53$, $p = .039$). *Post hoc* Scheffé tests revealed the difference on 'personal distress' to be between the AS and DPD groups (see figure 5.2) and the difference on 'perspective-taking' to be between the control and AS groups. No differences were found for the 'fantasy' items ($F_{(2, 42)} = 1.87$, $p > .05$) or 'empathic concern' ($F_{(2, 42)} = .71$, $p > .05$). These data are in keeping with the results from other measures of empathy, namely the EQ wherein group differences were found on 'cognitive empathy', and also the previous findings from these groups on the reduced 'personal distress' items from the IRI (see chapter 3). To summarise:

- Differences were found between the AS and DPD group on 'personal distress'.
- Differences were found between the AS and control group on 'perspective-taking'.

Figure 5.2 Mean and 95% confidence intervals for full score on IRI 'personal distress' items



Demographic factors

Analyses were performed to consider the role of demographic factors on the different components of empathy as measured by the IRI. Gender differences with women scoring higher were found on both the 'empathic concern' (t = -2.19, df 41, p = .034), and 'perspective-taking' (t = - 2.46, df 41, p = .018) – see Table 5.1. Differences were also found between high and low verbal IQ groups (based on a median split at 122) on the 'perspective-taking' scale (t = 1.99, df 40, p = 0.054). All these factors were therefore entered as covariates in the final analysis.

Full analysis

A factorial ANOVA was conducted with 'empathic concern' as the dependant variable and 2 between-participants factors being gender (male vs. female) and group (control/AS/DPD). There was no main effect for group ($F_{(2, 42)} = .04$, $p > .05$), gender ($F_{(1, 42)} = 3.01$, $p = .091$) nor an interaction between gender and group ($F_{(1, 42)} = .163$, $p > .05$). Lastly, an ANCOVA was conducted with 'perspective-taking' as the dependant variable, one fixed factor of group and a continuous covariate of verbal IQ. Again, there was no main effect for group ($F_{(2, 41)} = .89$, $p > .05$), or gender ($F_{(1, 41)} = 1.93$, $p > .05$), nor was there an interaction between gender and group ($F_{(2, 41)} = .50$, $p > .05$), but there was a significant main effect for verbal IQ ($F_{(1, 41)} = 6.36$, $p = .016$). This suggests the group difference previously observed on 'perspective-taking' may result from variation in IQ.

Table 5.1 Mean and SD scores for each group on IRI Subscales

		iri empathic concern				iri perspective taking			
		sex				sex			
		male		female		male		female	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
group	ctrl	17.78	2.73	19.73	5.04	16.56	2.35	18.67	5.27
	as	17.1	3.45	19.50	2.12	12.4	5.4	17	4.24
	dp	16.40	4.16	20.50	.71	16.2	7.69	19.50	.71

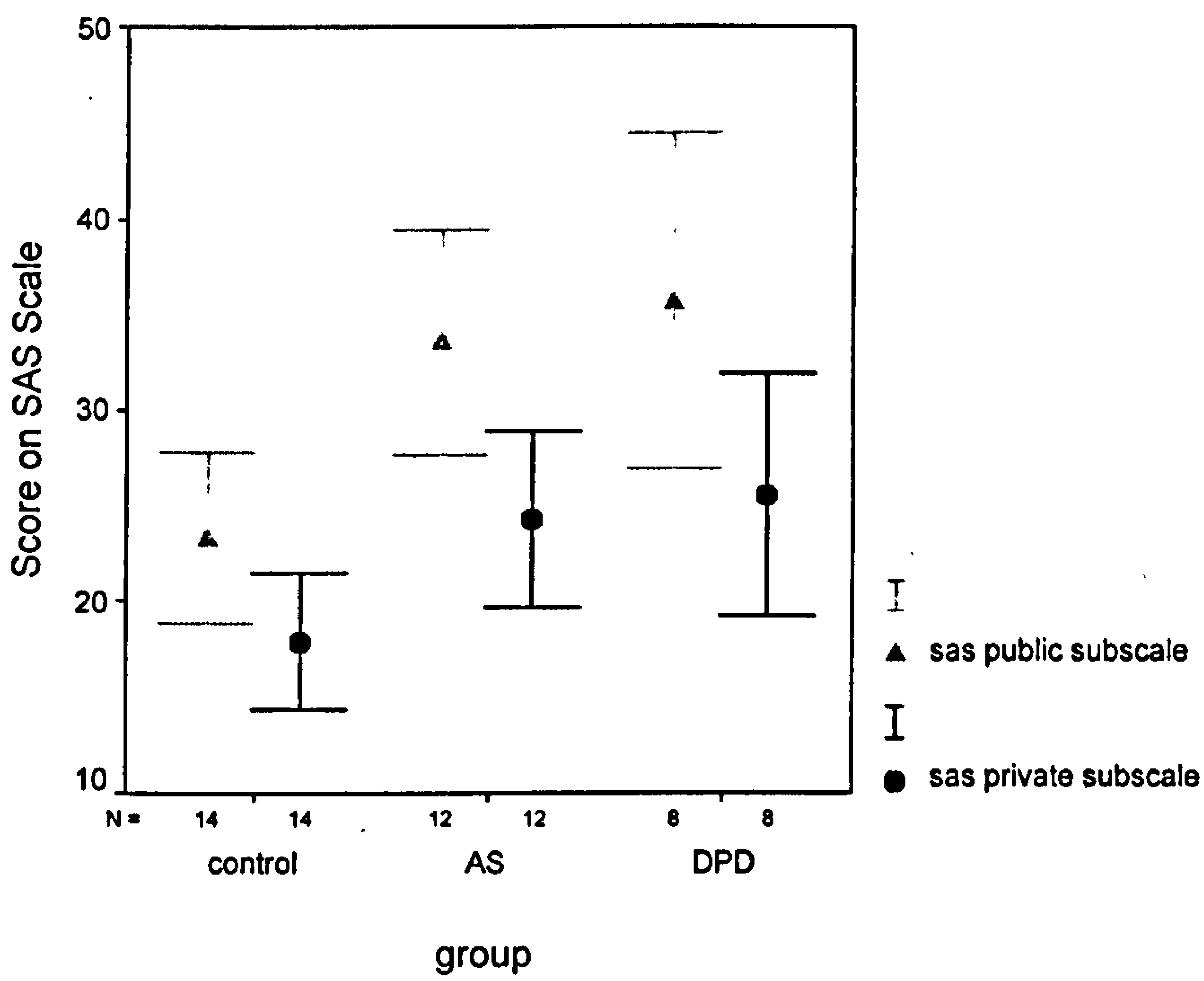
To summarise:

- No significant group differences were found for 'empathic concern' even with demographic factors held constant.
- No significant group differences were found for 'perspective-taking' and verbal IQ accounted for much of the variance.

Regarding the Self Absorption Scale (SAS) subscales, 2 one way ANOVAs were conducted for group and each subscale. There were main group effects for each subscale [public: $F_{(2, 29)} = 4.149$, $p = .027$/private: $F_{(2, 29)} = 4.12$, $p = .027$). *Post hoc*

Scheffé tests revealed there to be weak evidence of differences between the clinical and control groups on both the private (DPD: p = .092/AS p = .053) and public (DPD: p = .059/AS p = .074) self-absorption scales (see Figure 5.3). There were no demographic differences on either subscale of the SAS.

Figure 5.3: Mean scores with 95 % confidence intervals on self-absorption scale.



3.2 Affective state labelling (vignettes)

Simple group analyses

Although the data did not fully meet the parametric assumptions and were non-normally distributed, there was homogeneity of variance ($F_{(2, 77)}$ = .12, p > .05). A one-way ANOVA was therefore conducted to examine the effect of group membership on total number of correctly labelled affective states. No group differences were found for this variable ($F_{(2, 79)}$ = 2.12, p > .05) suggesting all groups could appreciate the emotional state of the protagonist.

Demographic Factors

In order to explore this lack of significant association further, the influence of demographic factors was also examined. Affective state attribution was not found to be related to age (t = -.32, df 76, p > .05), gender (t = -1.48, df 78, p > .05), verbal IQ (t = .36, df 76, p > .05), education (t = -.66, df 76, p > .05) or occupation (t = 1.37, df 78, p > .05).

Full analysis

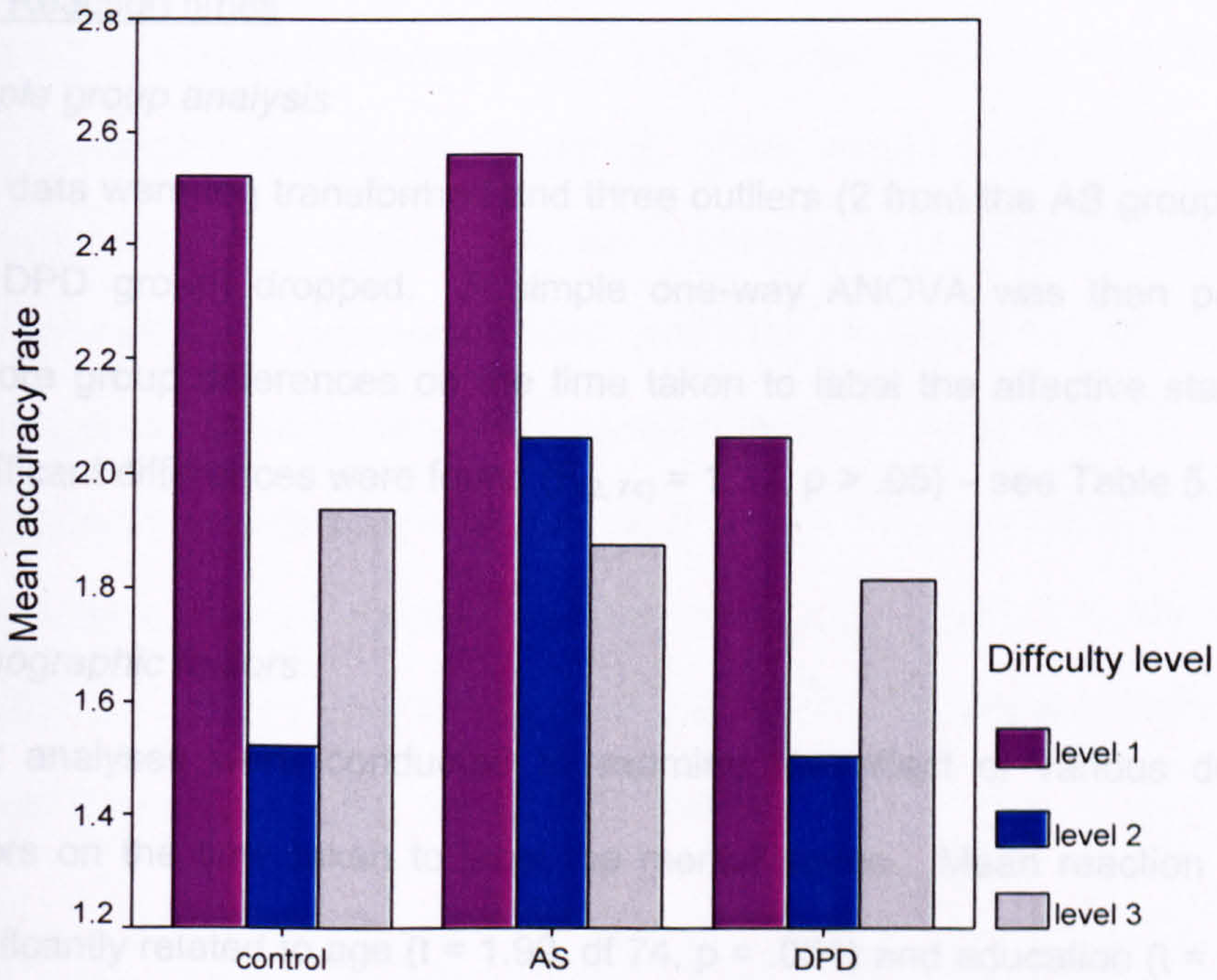
Subsequently, a 3 x 3 repeated measures ANOVA was performed to examined the relationship between the differing difficulty levels of emotion attribution in the 3 groups (see table 5.2) in an attempt to tease apart the processes underlying affective state attribution.

Table 5.2: Mean and SD score on items of different difficulty levels

		total score on level 1 questions		total score on level 2 questions		total score on level 3 questions	
		Mean	SD	Mean	SD	Mean	SD
group	CONTROL	2.52	.58	1.52	.95	1.94	.89
	AS	2.56	.51	2.06	.77	1.88	.81
	DPD	2.06	1	1.5	.73	1.81	.75
Total		2.44	.69	1.63	.89	1.9	.84

The dependent variable was difficulty (with 3 levels) and the independent variable group (AS/DPD/control). The assumption of sphericity was met and there was homogeneity of variance on all but level 1. A main effect was found for 'difficulty levels' [F (2, 154) = 13.85, p = .001] and paired t tests revealed the differences to be between all the levels (level 1 vs. 3: t = 4.91, df 79, p = .001/level 1 vs. 2: t = 6.4, df 79, p = .001/ level 2 vs. 3: t = 2.43, df 79, p = .028). However, as with the control data in chapter 4, participants tended to find level 2 harder than level 3 regardless of group membership (see figure 3). Although, despite the lack of a significant interaction between group and 'difficulty level' ($F_{(4, 154)} = 1.54$, $p > .001$) or the main effect for group ($F_{(2, 77)} = 2.12$, $p > .05$), the plot did suggest that the AS group tended to find level 2 easier than the other groups (see figure 5.4) and that the DPD group found level 1 more difficult. This may suggest the use of different strategies by each group.

Figure 5.4: Mean accuracy rate on affective state labelling task



There was no difference between groups on the non-emotional vignette ($F_{(2, 79)} = 1.45$, $p > .05$).

In summary:

- There were no significant group differences on the affective state labelling task even with demographic factors taken into consideration.
- Accuracy rates differed between 'difficulty' levels with participants finding 'level 2' the most difficult.

3.3. Reaction times

Simple group analysis

The data were log transformed and three outliers (2 from the AS group and 1 from the DPD group) dropped. A simple one-way ANOVA was then performed to explore group differences on the time taken to label the affective states, and no significant differences were found ($F_{(2, 74)} = 1.37$, $p > .05$) – see Table 5.3.

Demographic factors

Next analyses were conducted to examine the effect of various demographic factors on the time taken to label the mental states. Mean reaction times were significantly related to age ($t = 1.99$, df 74, $p = .051$) and education ($t = 3.34$, df 76, $p < .001$). However, they were not related to gender ($t = .6$, df 76, $p > .05$), verbal, IQ ($t = -1.58$, df 74, $p > .05$) or occupation ($t = -1.26$, df 76, $p > .05$). Age and education were therefore included in subsequent analyses which examined reaction times for each vignette.

Table 5.3: Mean and SD reaction times on each vignette

		Vignette B		Vignette C		Vignette D	
		Mean	SD	Mean	SD	Mean	SD
group	CONTROL	1.43	.17	1.33	.14	1.38	.13
	AS	1.54	.21	1.40	.17	1.46	.13
	DPD	1.47	.13	1.35	.15	1.40	.15
Total		1.46	.17	1.35	.15	1.40	.14

* NB these data have been normalised using a logarithmic transformation

Full analysis

A 3 x 3 x 2 repeated measures ANOVA was then performed with reaction time as the dependant variable (3 vignettes), group with 3 levels, age co-varied and education entered as an additional between-group factor. As there was homogeneity of regression lines and the assumptions of sphericity and homogeneity of variance were met, a full factorial model was run. There was a significant main effect for reaction time ($F_{(2, 134)} = 3.65$, $p = 0.028$) but no significant interactions between reaction time and age ($F_{(2, 134)} = .451$, $p > .05$), group ($F_{(4, 134)} = .068$, $p > .05$) or higher education ($F_{(2, 134)} = 1.36$, $p > .05$) and neither was there a 3 way interaction ($F_{(4, 134)}$ .21, $p >.05$) - see Table.5.4.

Table 5.4: Within-participants effects for reaction times from repeated measure ANOVA

Tests of Within- Participants Effects

	Type III Sum of Squares	df	Mean Square	F	Sig.
Reaction Time	6.920E-02	2	3.460E-02	3.665	.028
Reaction Time * AGE	8.515E-03	2	4.257E-03	.451	.638
Reaction Time * Group	2.562E-03	4	6.406E-04	.068	.991
Reaction Time * Education	2.572E-02	2	1.286E-02	1.362	.260
Reaction Time * Group * Age	7.907E-03	4	1.977E-03	.209	.933
Error (Reaction Time)	1.265	134	9.440E-03		

Paired t-tests revealed significant differences between each vignette with participants taking longest to attribute affective states in vignette b followed by d and c (b vs. c: t = 7.11, df 77, p <.001/b vs. d t = 3.9, df 79, p = .0011/c vs. d: t = -.3.34, df 77, p = .01). There was no main effect for group on reaction time (see table 5.5) perhaps suggesting that all groups used similar strategies ($F_{(2, 67)}$ = .503, p > .05).

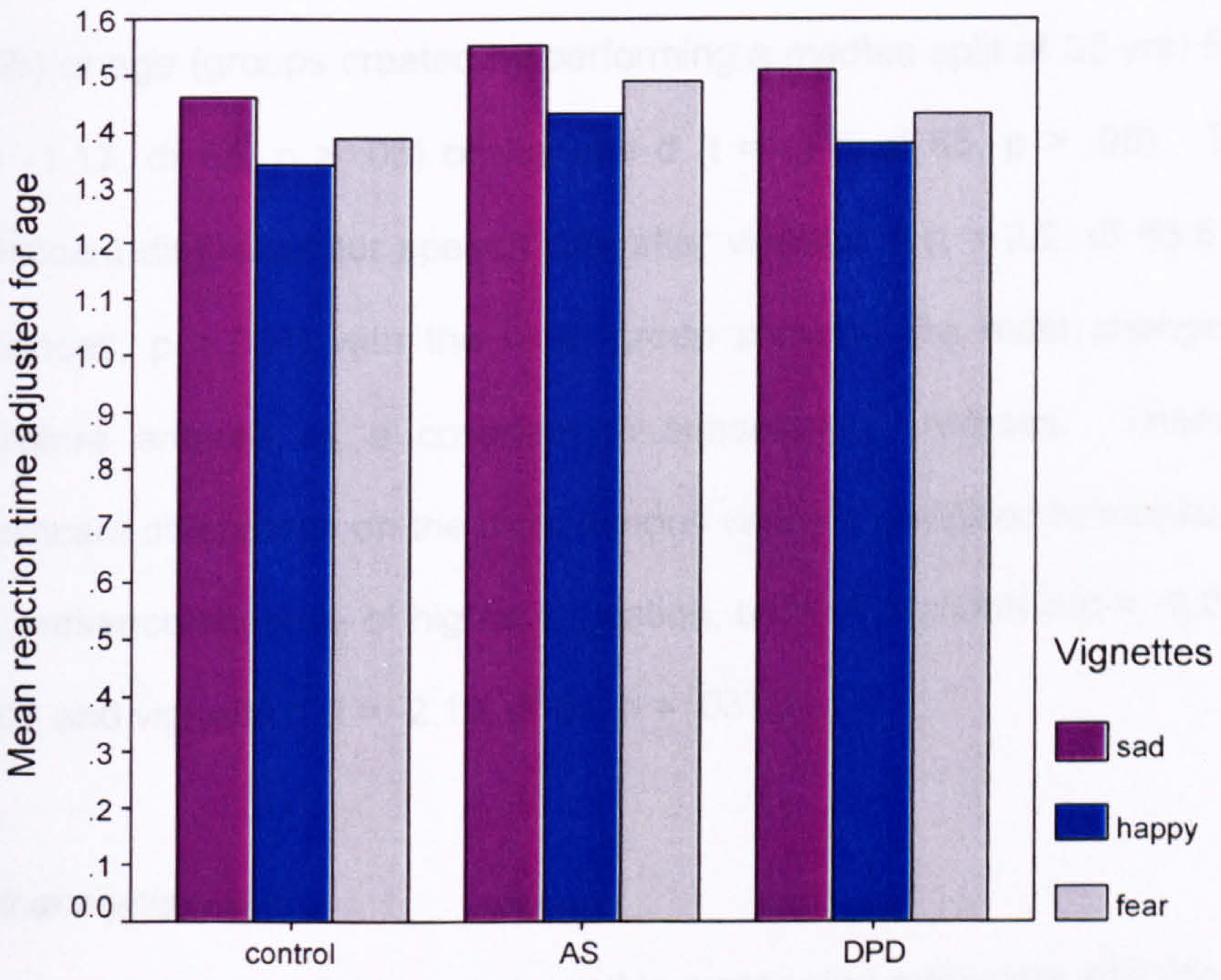
Table 5.5: Between-participants effects for reaction time with age co-varied from repeated measures ANOVA

Tests of Between-Participants Effects

	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	27.909	1	27.909	581.825	.001
AGE	.145	1	.145	3.027	.086
GROUP	4.825E-02	2	2.413E-02	.503	.607
HIGHER	.150	1	.150	3.123	.082
GROUP * HIGHER	1.498E-02	2	7.488E-03	.156	.856
Error	3.214	67	4.797E-02		

The plot based on the estimated means after adjusting for age (see Figure 5.5) indicate that the AS group did react slightly slower overall in the task but this was not statistically significant. That this group showed the same pattern as the other groups suggests that all groups adopted the same kind of strategy.

Figure 5.5: Mean reaction times on labelling task adjusted for age



To summarise:

- There were no significant differences between groups in the time taken to label the affective states.
- Age was associated with reaction times on this task.
- There were reaction time differences between vignettes.

3.4. Speech rate

Simple analyses were not conducted on this variable as summing the speech rate data from each vignette, would not be meaningful.

Demographic factors

Parametric tests were conducted to explore the effects of demographic variables on a log transformation of speech rate. There was no effect for gender (vignette b: $t = -.85$, df 68, $p > .05$ /vignette c: $t = .91$, df 67, $p > .05$/vignette d: $t = .78$, df 67, $p > .05$) or age (groups created by performing a median split at 32 yrs) for vignette c ($t = -1.17$, df 65, $p > .05$) or vignette d ($t = -.27$, df 65, $p > .05$). There was a significant difference for speech rate after vignette b ($t = 2.2$, df 63.5 for unequal variances, $p = .031$) with the older group showing the most change. Age was therefore entered as a covariate in subsequent analyses. There were also significant differences on the dichotomous variable designed to measure education i.e. presence/absence of higher education, both on vignette b ($t = -2.06$, df 66, $p = .043$) and vignette d ($t = -2.13$, df 65, $p = .037$).

Full analysis

Log-transformed variables were used in a repeated measures ANOVA with speech rate as the within-participant variable with 3 levels (sad/happy/fear) and 2 between group factors: education with (presence and absence of higher education) and group (control/AS/DPD).

Due to technical problems, 5 people from the AS group and 3 people from the DPD group had the speech rate dataset missing. In addition, 2 x higher education values were missing - and the age data for 1 person in the ~~AS~~ control group. To maximise the degrees of freedom in the repeated measures ANOVA, the missing age data were replaced with the series mean and the education data replaced with one value from each category.

As there was homogeneity of regression slopes for age and group, and sphericity, a full factorial model was performed. The main effect for speech rate approached significance ($F_{(2, 61)}$ = 2.83, p = .063). There was also a significant interaction between speech rate and age ($F_{(2, 122)}$ = 3.9, p = .023) and between speech and group ($F_{(4, 122)}$ = 2.4, p = .054) suggesting these factors had an effect on speech rate for some of the vignettes (see Table 5.6).

Table 5.6: Within-participants effects for speech rate from repeated measures ANOVA

Tests of Within- Participants Effects

	Type III Sum of Squares	df	Mean Square	F	Sig.
Speech Rate	5.845E-02	2	2.922E-02	2.825	.063
Speech Rate * Age	8.064E-02	2	4.032E-02	3.898	.023
Speech Rate * Education	1.509E-02	2	7.547E-03	.730	.484
Speech Rate * Group	9.914E-02	4	2.479E-02	2.396	.054
Speech Rate * Education * Group	4.408E-02	4	1.102E-02	1.066	.377
Error (Speech)	1.262	122	1.034E-02		

Regarding between-participants effects, there were no main effects for group ($F_{(2, 58)}$ = .54, p > .05) or education (see table 5.7) although again this approached significance ($F_{(1, 61)}$ = 3.07, p = .085). However, there was no interaction between speech rate and education ($F_{(2, 122)}$ .73, p > .05), speech rate, education and age ($F_{(4, 122)}$ = 1.07, p > .05) or any interaction between group and higher education ($F_{(2, 61)}$ = 1.84, p > .05).

Table 5.7: Between-participants effects for speech rate with age as a covariate from repeated measures ANOVA

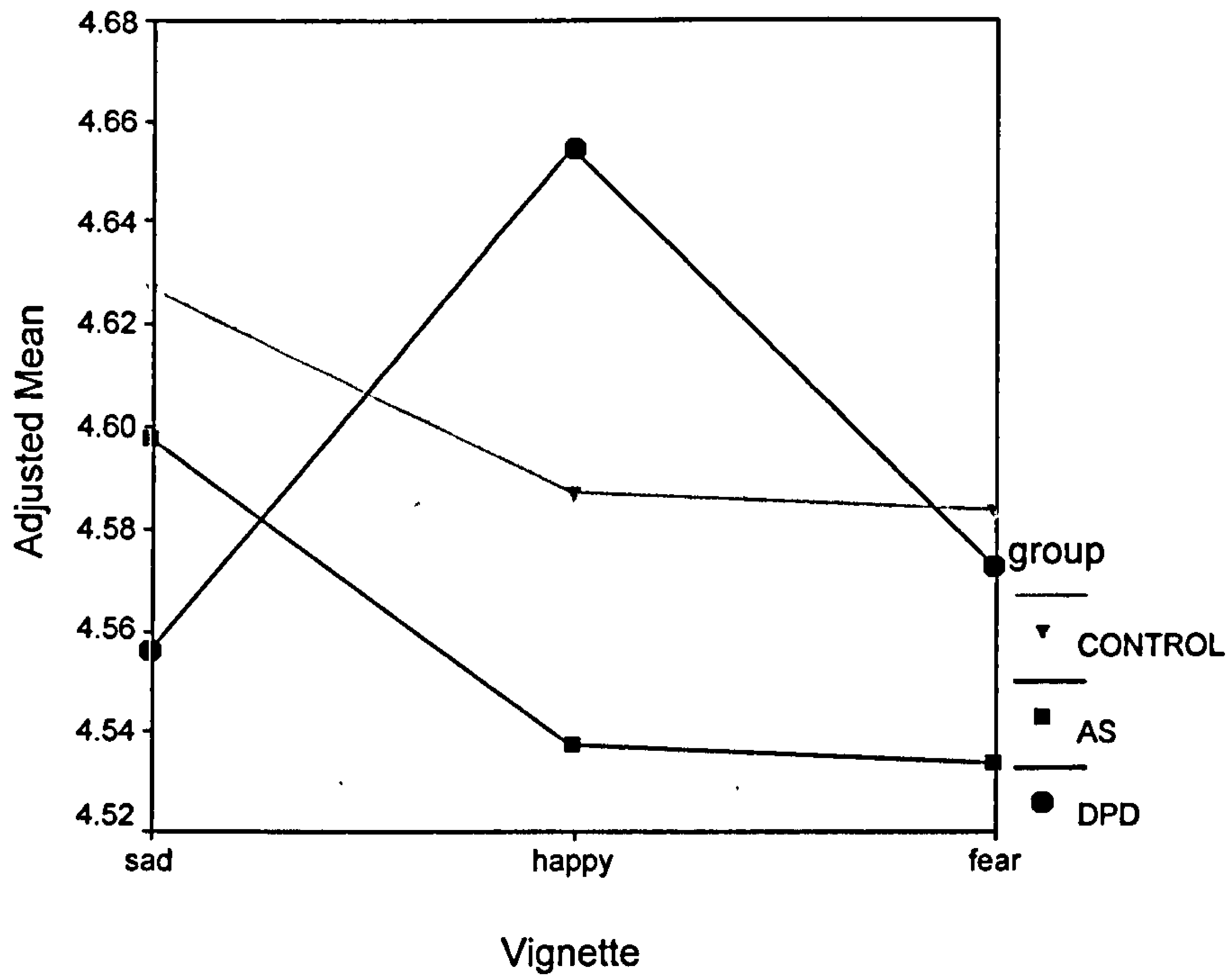
Tests of Between-Participants Effects

	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	105.912	1	105.912	8320.534	.001
AGE	3.271E-07	1	3.271E-07	.000	.996
EDUCATION	3.912E-02	1	3.912E-02	3.074	.085
GROUP	1.376E-02	2	6.882E-03	.541	.585
EDUCATION * GROUP	4.683E-02	2	2.341E-02	1.839	.168
Error	.776	61	1.273E-02		

Planned contrasts revealed the near significant speech rate and group interaction to be between the 'sad' and 'happy' vignettes ($F_{(2, 61)} = 4.28$, $p = .021$). Examination of the plot of means adjusted for age, showed that the DPD group performed significantly different to the other groups (see Figure 5.6 below). Paradoxically, it appears the DPD participants decelerated their speech after the happy vignette while the other groups tended to accelerate and there was also a tendency for the DPD group to accelerate after the sad vignette again in contrast to the other groups. This is indicative of a lack of congruent emotional response, as such congruencies are taken to be characteristic of parallel empathy. In summary:

- Speech rate differed between vignettes in the expected directions.
- Age was associated with speech rate with older people slowing down more after the sad vignette.
- The DPD group displayed differential speech rate to the other groups.

Figure 5.6: Mean speech rate change adjusted for age



* Original data were expressed as a percentage of the baseline so an increase on y axis represents deceleration and decrease represents acceleration from the baseline, following each vignette.

3.5 Trait Task

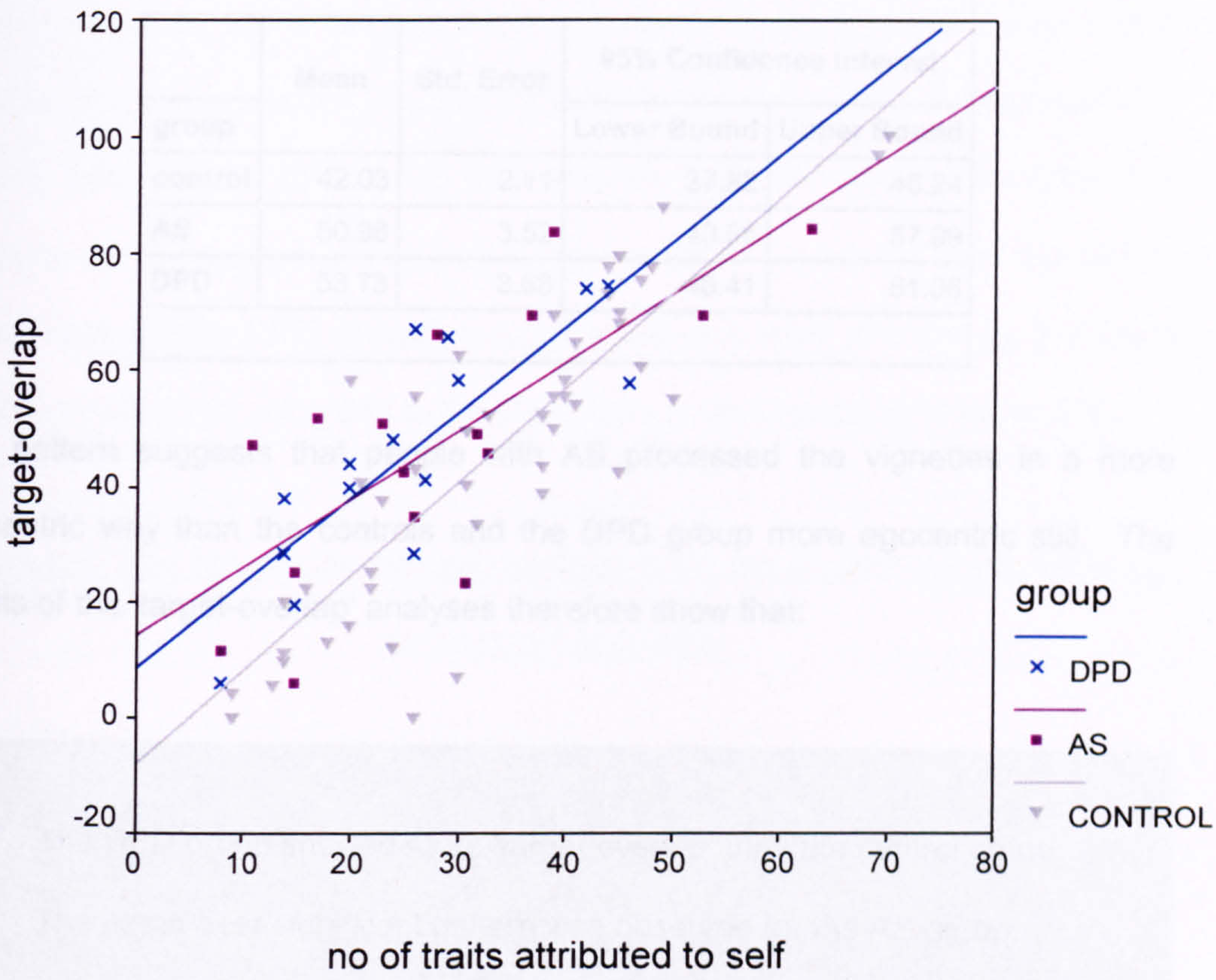
There were two dependant variables; i) self-overlap variable - percentage of the 'self' traits that were later endorsed for the target, and ii) target-overlap variable - percentage of traits attributed to the target that were earlier endorsed for the self. The self-overlap variable was normalised using a square root transformation. There are also two further variables key to these analyses: i) 'mean no of traits attributed to the target overall' and ii) 'number of traits attributed to self' (see below).

Consistent with the original study (Davis *et al.*, 1996), the association between these two additional variables of interest i) 'mean no of traits attributed to the target overall' and ii) 'no of traits attributed to self', and the dependant variables was explored from the outset. Both these variables were predictors of trait overlap in chapter 4. Although, there were no group differences on 'mean number of traits attributed to the target' ($F_{(2, 77)} = 1.07$, $p > .05$) or 'number of traits attributed to self' ($F_{(2, 77)} = 2.02$, $p > .05$) there were significant associations within-groups for both variables for all 3 groups. Therefore, the decision was made to enter these measures as covariates in all analyses. However, as both indices of number of traits were also inter-correlated, only the measure with the strongest correlation with the dependant variable was entered into each analysis – 'number of traits attributed to self' in the target-overlap analysis and 'mean number of traits attributed to the target' in the self-overlap analysis. These were also the significant predictors in chapter 4. These variables were therefore included from the outset

Group analyses

There was homogeneity of the regression lines (see Figure 5.7) and so an ANCOVA controlling for 'number of self traits' was conducted for group and target-overlap. There was a main effect for the covariate 'number of self traits' as expected ($F_{(1, 77)} = 168.08$, $p > .001$). There was also a main effect for group ($F_{(2, 77)}$ $= 5.1$, $p = .008$).

Figure 5.7: Mean % target-overlap scores and number of traits attributed to self



Post hoc Scheffé tests revealed the difference to be between the control and DPD groups (p = .019) with the difference between the AS and control group approaching significance (p = .089). Examination of the means show that both clinical groups had a tendency to show more target-overlap (see Table 5.8).

Table 5.8: Mean % target-overlap score adjusted for 'number of self traits endorsed

group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
control	42.03	2.11	37.82	46.24
AS	50.98	3.52	43.96	57.99
DPD	53.73	3.68	46.41	61.06

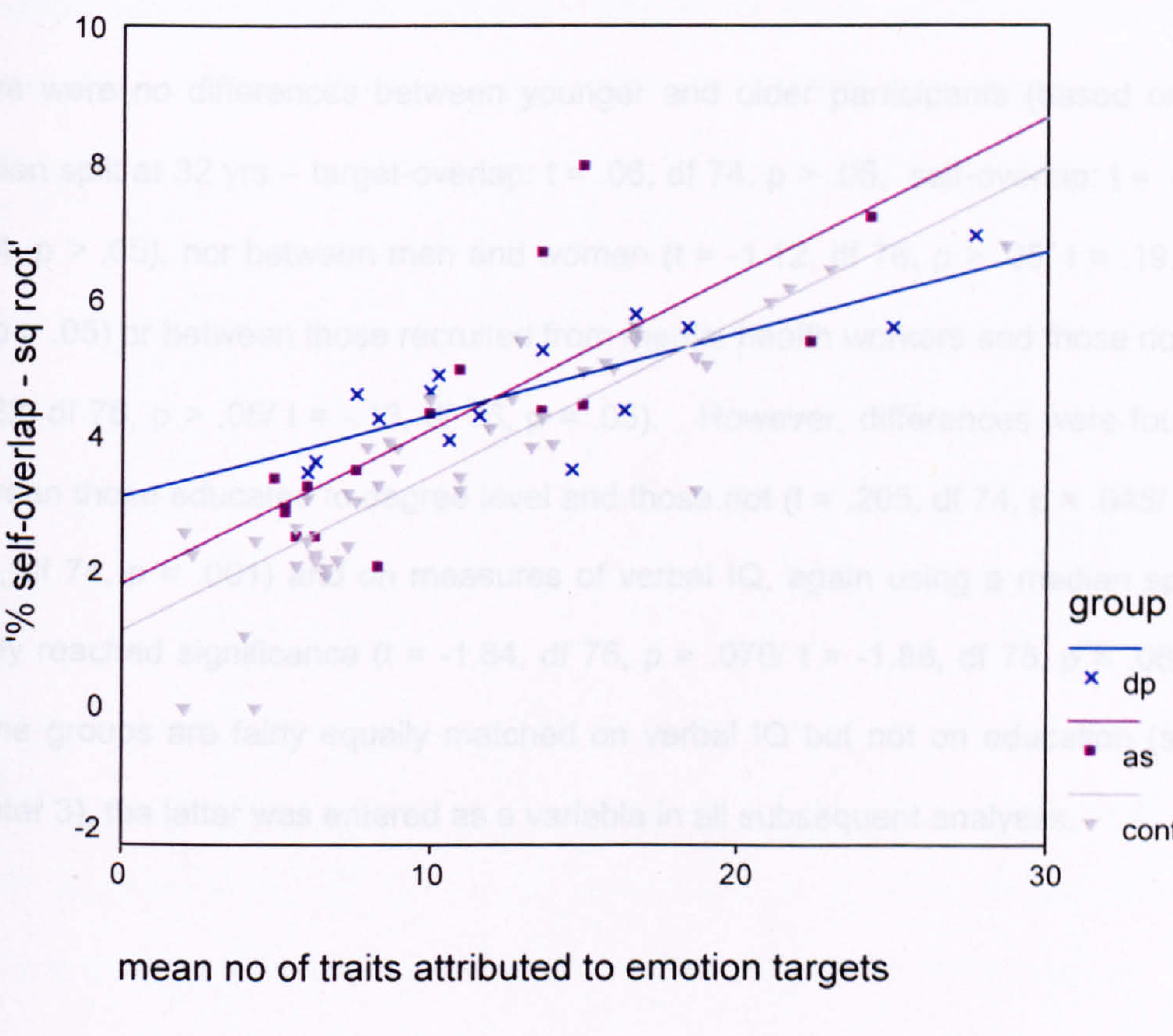
This pattern suggests that people with AS processed the vignettes in a more egocentric way than the controls and the DPD group more egocentric still. The results of the 'target-overlap' analyses therefore show that:

- The DPD group showed more 'target-overlap' than the control group.
- The same near significant pattern was observed for the AS group.

The same type of analysis was also performed to explore the association between group membership and 'self-overlap' (as opposed target-overlap in the previous section). Based on previous findings (see above and chapter 4) the means were adjusted to reflect 'mean number of traits attributed to the target'. In this instance, the regression lines were heterogeneous between groups ($F_{(2, 77)} = 4.03$, $p = .022$) having important consequences for interpretation of the data (see figure 7). With this interaction built into the model there was also a main effect for group ($F_{(2, 77)} = 6.27$, $p = 0.003$) and for the covariate 'mean number of traits attributed to the target' after all the affective state vignettes ($F_{(1, 77)}$ 164.73, $p < .001$).

Post hoc Scheffé tests revealed significant differences between the AS and control group (p = .022) with the difference between the DPD and control group approaching significance (p = .087). This suggests that once again the clinical groups displayed more overlap than controls. However, as the relationship between the 'mean number of traits attributed to the target' and 'self-overlap' was not the same for the DPD group as it was for the other groups, this difference needs to be considered when interpreting the data. It seems from the plot of means that for the DPD groups' overlap score was least affected by the gross number of traits they attributed to the characters in the vignette (see figure 5.8).

Figure 5.8: The association between target-overlap, the number of traits attributed to the target and group



To summarise:

- The AS group displayed more overlap, as measured by the 'self-overlap' variable, than controls.
- A similar, near significant pattern was observed between those with DPD and controls.

Demographic factors

In order to explore the data further, we looked at age, gender, whether or not the participant entered higher education, whether or not their occupation was mental health worker at the IOP and verbal IQ as estimated from the NART on i) the self-overlap variable i.e. percentage of the 'self' traits that were later endorsed for the target and ii) target-overlap variable, that is, percentage of traits attributed to the target that were earlier endorsed for the self.

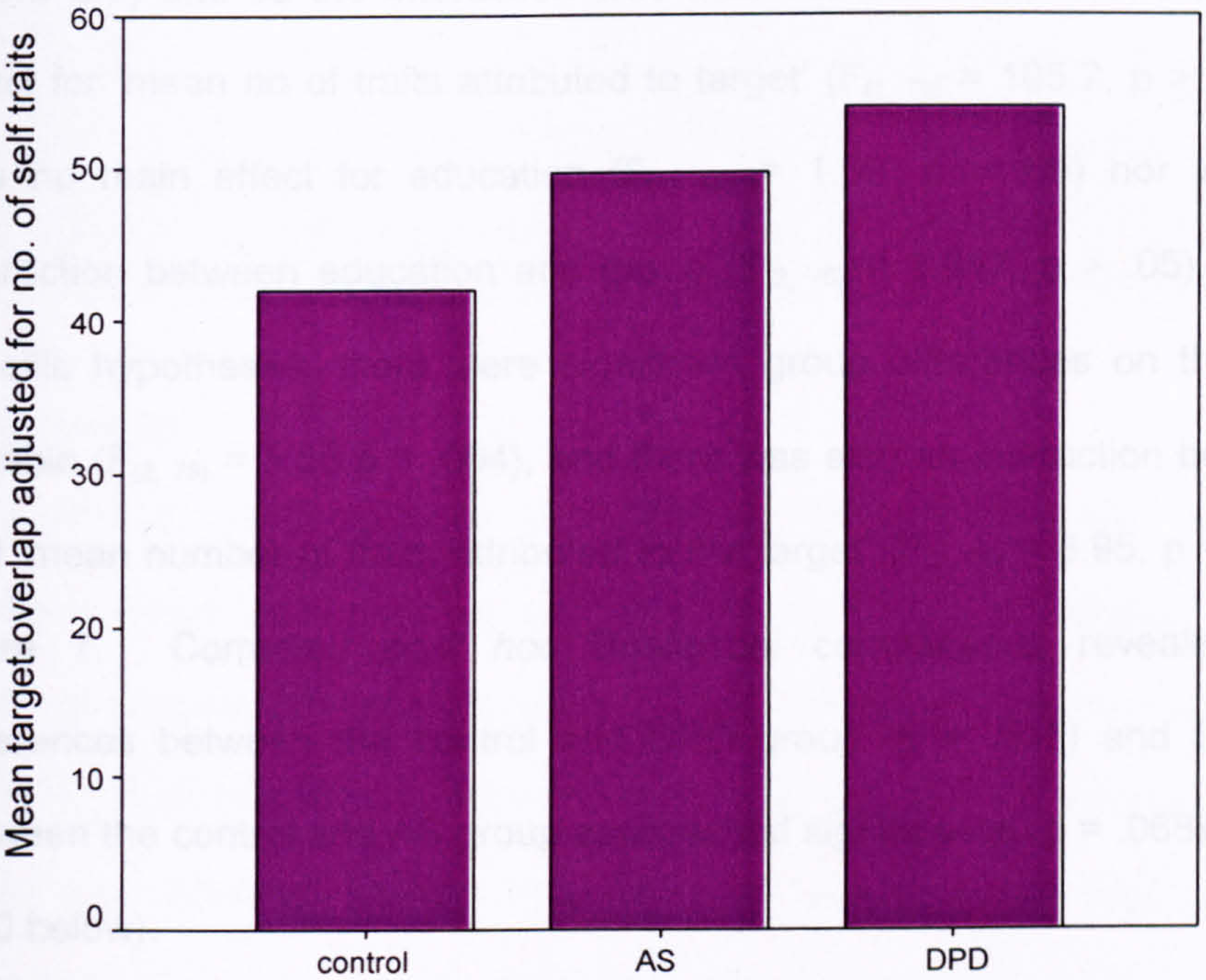
There were no differences between younger and older participants (based on a median split at 32 yrs – target-overlap: $t = .06$, df 74, $p > .05$, self-overlap: $t = .62$, df 74, $p > .05$), nor between men and women ($t = -1.12$, df 76, $p > .05$/ $t = .19$, df 76, $p > .05$) or between those recruited from mental health workers and those not ($t = -.23$, df 76, $p > .05$/ $t = -.13$, df 76, $p = .05$). However, differences were found between those educated to degree level and those not ($t = .205$, df 74, $p = .045$/ $t = 3.43$, df 74, $p = .001$) and on measures of verbal IQ, again using a median split, nearly reached significance ($t = -1.84$, df 75, $p = .070$/ $t = -1.88$, df 75, $p = .063$). As the groups are fairly equally matched on verbal IQ but not on education (see chapter 3), the latter was entered as a variable in all subsequent analyses.

- The only demographic factors associated with 'self-overlap' and 'target-overlap' were education and verbal IQ.

Full analysis

A factorial ANCOVA with 2 between group factors – 'group' with 3 levels and education with 2 levels – and 1 covariate – number of traits attributed to 'self', was performed with the dependant variable as 'target' overlap. The dataset conformed to parametric and other assumptions. The covariate (no. of traits attributed to the self) showed a significant ($F_{(1, 75)} = 158.6$, $p < .001$) association with the 'target-overlap' as expected. There was no main effect for education ($F_{(1, 75)} = 1.96$, $p > .05$) and neither was there an interaction between education and group ($F_{(2, 75)} = 2.76$, $p > .05$). Regarding the specific hypotheses, there were significant group differences on the dependant variable ($F_{(2, 75)} = 3.86$, $p = .026$) – see figure 8. Corrected *post hoc* Bonferroni adjusted comparisons reveal significant differences between the control and DPD group ($p = .025$) with the DPD group displaying more target-overlap.

Figure 5.9: Mean % target-overlap adjusted for no. of self traits

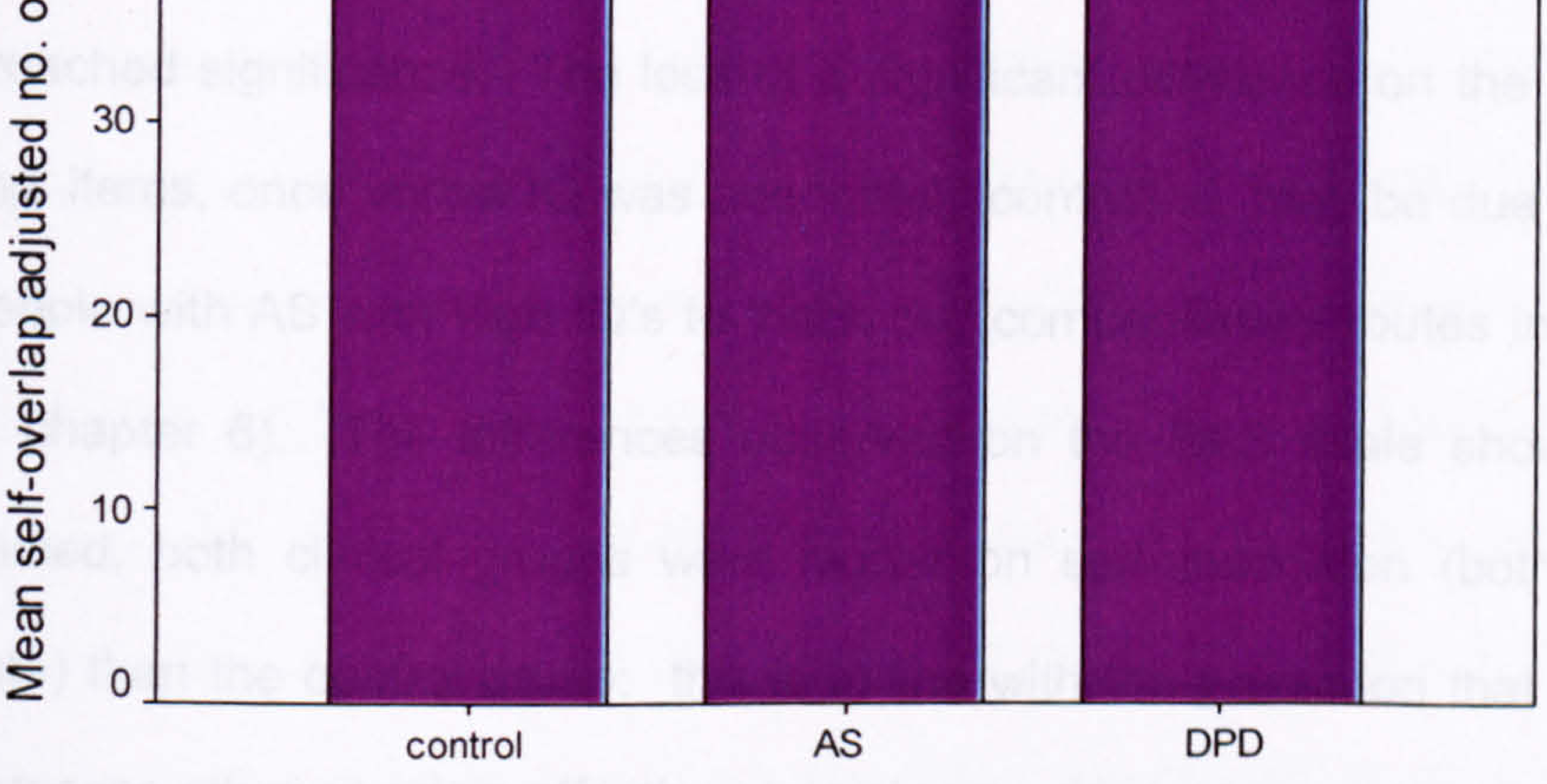


It seems therefore that:

- Compared to controls, the DPD group showed more overlap as measured by the 'target-overlap' variable, even with demographic factors considered.
- The similar pattern previously observed for the AS group on this variable disappeared once the relevant demographic factors were statistically controlled.

The same analysis was also conducted for the alternative index of 'self-overlap' with 'mean number of traits attributed to the targets' as a covariate. With the additional grouping variable the data still conformed to the homogeneity of variance principle ($F_{(5, 70)} = 2.18$, $p > .05$). Again there was an interaction between the covariate and group, suggesting a lack of homogeneity of regression slopes (see Figure 5.9) and so the interaction was built into the model. There was a main effect for 'mean no of traits attributed to target' ($F_{(1, 75)}$ = 105.7, p > .001). There was no main effect for education ($F_{(1, 75)}$ = 1.59, p > .05) nor was there an interaction between education and group ($F_{(2, 75)}$ = 1.842, p > .05). Regarding specific hypotheses, there were significant group differences on the dependant variable ($F_{(2, 75)}$ = 5.86 p = .004), and there was also an interaction between group and 'mean number of traits attributed to the target' ($F_{(2, 75)}$ = 3.95, p = .024) – see figure 7. Corrected *post hoc* Bonferroni comparisons revealed significant differences between the control and DPD group (p = .015) and the difference between the control and AS group approached significance (p = .068) – see Figure 5.10 below).

Figure 5.10: Mean % self- overlap adjusted for no. of traits attributed to target



- The AS group show a near significant tendency to display more 'self-overlap' than the control group.
- The DPD group show more 'self-overlap', however, they also show a differential relationship between the dependant variable and covariate (see below for further discussion).

4. Discussion

As predicted, all the groups were able to label the affective states of the protagonist and there were no reaction time differences, suggesting they all used similar strategies. The clinical groups showed significantly more trait overlap than controls which may be indicative of egocentric simulation. In addition, the DPD group also showed speech rate patterns, suggesting a lack of congruent physiological arousal. This is in keeping with reports from people who suffer from DPD of a subjective emotional blunting.

In line with the data from chapter 3, there were significant differences between the DPD and AS groups on the full scale IRI 'personal distress' items. The failure to detect a difference between the control and clinical groups may have been due to the small sample size, although the difference between the AS and control groups approached significance. The loss of a significant difference on the 'perspective-taking' items, once verbal IQ was adequately controlled, may be due to the ability of people with AS with high IQ's to 'hack out' compensatory routes in this domain (see chapter 6). The differences observed on the SAS scale showed that, as predicted, both clinical groups were higher on self-absorption (both public and private) than the control group; this is in line with the prediction that both groups will process other people's affective states in an egocentric way (see below).

Findings from the mental state labelling task show that all participants could label the non-affective epistemic states in the control vignette, confirming that people with AS can appreciate some simple representational states. On the affective state vignettes, an effect for difficulty level was also found, indicating that this manipulation was partially successful. However, all groups generally found level 2 more difficulty than level 3 which might be explained by previous findings that suggest people tend to pick up broad rather than specific affective states from written narratives (Gygax *et al.*, 2003). This may be especially so when one considers that level 2 emotional states were designed to be somewhat idiosyncratic and/or unusual reactions to the situation. In replication, a larger range of scores may be more sensitive to individual variation.

The reaction times in each group were similar, suggesting that all the groups may have used the same strategy. Although the AS group were slightly slower on all the vignettes this was not statistically significant; this is of particular interest as some previous studies have suggested that people with AS may use effortful, rule-based strategies which would be expected to manifest a longer reaction time. Although an alternative explanation is that the task prompted all groups to use a rule-based strategy. This would mean that the lack of group differences cannot be interpreted safely to suggest the AS group fully appreciated the affective states displayed by the protagonist. Zahavi and Parnas (2003) point out that the common strategy advocated by 'theory-theorists' for appreciating others' mental states is remarkably similar to that understood to be the basis of compensatory strategies in people with ASDs and high IQ's i.e. 'wooden algorithms and formulae' (see also Belmonte and Yurgelun-Todd, 2003). Consistent with this idea, there is also the possibility that each difficulty level evoked a different strategy, which may have confounded the reaction time data. The AS group tended to do *better* than both other groups on level 2, the hardest level, although again this result was not significant. This may indicate that these items required a strategy at which the AS group were more skilled such as rule-based processing based on such 'algorithms and formulae'.

Differential processing was also evident on 'level 1' items of the task. There was a tendency for the DPD group to endorse more incorrect basic emotional reactions and examination of the raw data suggest that this was the case for a handful of participants in this group, explaining the heterogeneity of variance for this particular difficulty level. One explanation is that the comprehension of simple basic emotions may be easier if the qualitative dimensions can be appreciated, and perhaps a covert online somatosensory representation constructed (see Adolphs et al., 2000). Some people with DPD may find this difficult and indirect evidence of this is provided by the speech rate data (see below). Furthermore, people with DPD have been shown not to activate some of the brain areas involved in constructing implicit online somatosensory representations (Phillips *et al.*, 2001). However, despite the possible different processing styles adopted for each level of difficulty, none of these effects were statistically significant and overall both clinical ~~control~~ groups performed in a manner similar to controls. Bearing this in mind it seems that these data need to be considered in the light of the other components of this task, physiological arousal and the use of the self-concept.

In all the groups, physiological arousal as measured by speech rate after the sad vignette was associated with age, with older people displaying more enhanced mood-congruent speech rate (they slowed down more). However, even after controlling for age, an interaction remained between speech rate and group with the DPD group displaying differential speech rate on the sad and happy vignettes. The direction of this difference indicates a lack of congruent physiological arousal as they decelerated after the happy vignette and accelerated after the sad vignette, in contrast to the other groups. This suggests a lack of congruent emotional response to other people's affective states. However, this is contrary to the data from the self-report measures (see chapter 2 and 3) which suggest that people with DPD do feel emotional empathy and empathic concern. One explanation is that this discrepancy may be due to the items on the self-report measures tapping more than simply physiological reactions, such as meta-representational ability.

Interestingly, the DPD group still showed some physiological reactions to the stimuli i.e. they did not stay at baseline, but their responses were inconsistent with an empathic response. This may suggest that the 'emotional blunting' they describe may be the subjective result of an 'emotional mismatch'. Lastly, there were no group differences on the control or fear vignettes. This finding may be due to the fact that speech rate measures do not distinguish between 'personal distress' and 'parallel empathy'. However, this group's IRI 'personal distress' scores were low suggesting they are not prone to this state, although it is true to say their scores did not represent a complete absence of distress in response to displays of emotion.

As predicted, the AS group performed remarkably similarly to the control group on this task. Although their speech rate appeared to be generally faster ~~slower~~ than the other groups this was not significant (see figure 5). This adds to the now growing evidence that people with AS (or indeed ASDs) can respond emotionally to other people's affective states inasmuch as they have the ability to experience affective empathy. Consistent with this, no group differences were found on the 'empathic concern' items on the IRI. The AS group also reported high levels of 'personal distress' and showed physiological arousal in accordance with parallel affective empathy e.g. in the 'sad' vignette. This may shed light on the interrelation between distress and arousal, suggesting that affective empathy may be a precursor to personal distress (see Figure 1). It seems that the idea that people with ASDs are 'emotionally cold' may be misguided, or at least an over-generalisation, and instead their problems lie in other domains of social cognition.

In terms of the task, there were distinctive speech rate patterns for each of the affective states as predicted. However, the dataset as a whole needs to be interpreted with caution because of lack of power due to missing data. Furthermore, as discussed in chapter 4 this measure has neither the sensitivity to pick up reactive states such as sympathy and compassion nor to distinguish between egocentric and empathic reactions.

The data from the trait task suggests a tendency for both clinical groups to process others' affective states egocentrically, in line with their scores on the self-absorption scale (see above). The DPD group scored consistently higher than the control group on the trait overlap scores, suggesting a more extreme reliance on the self-concept when attributing affective states. For both types of trait overlap variable, differences were found between control and DPD groups, over and above the effects of the covariate and demographic factors, suggesting the use of different cognitive processes. The AS group also showed significantly more overlap than the control group on the 'self-overlap' variable, although the near significant difference for the 'target-overlap' variable disappeared once demographic factors were controlled.

More specifically, for the 'self-overlap' variable, in addition to group differences, there was an interaction between sheer number of traits attributed to the target and group membership, indicating that the data deviates from the assumption of homogeneity of regression slopes. However, examination of the scatterplot (see Figure 6) suggests this interaction results from the DPD group showing *less* of an association between self-overlap and the covariate i.e. raw number of traits attributed to the target, than the other groups, consistent with the idea of the 'self' being the centre of the DPD sufferers' world. Furthermore, an effect of group still remained above and beyond the interaction with the covariate suggesting that some of the explained variance is independent of this interaction. *Post hoc* tests with adjusted significance levels show this additional variance to be due to significant differences between the control and AS groups. These tests also revealed a significant difference between the DPD and control group, but as explained, this has to be interpreted along with the interaction.

In terms of interpretation, it seems that the DPD group utilise the self-concept in reasoning about others' mental states, to a greater degree than people in the control group. These effects are independent of education and the 'number of traits attributed to the self'. As explained, the findings from the 'self' overlap variable may be in part due to the DPD group displaying self-other overlap which is to some extent independent from the actual number of traits attributed. To find that those with DPD rely more on the self-concept when decoding others' affective states, in contrast to those without DPD, is potentially important for devising clinical interventions. Results from previous chapters do not suggest that these groups are more empathic than the control group, and this would be contrary to clinical reports, ruling out a simple linear relationship between trait overlap and empathy. It seems much more likely that the relationship is more complex, with an optimum level such as that displayed by controls, with overlap over and above this indicative of differential processing. An excess of overlap may be best explained with reference to the 'egocentric/allocentric' distinction with additional overlap representing 'egocentric' simulation (see chapter 4 and Decety and Sommerville, 2003, Langdon and Coltheart, 2001). This would fit well with the newly emerging empirical literature on 'self-absorption' in DPD and provides support for therapeutic interventions that have a component geared towards defocusing the client from their own self-concept. It may well be that such a strategy, if successful, would have the indirect result of reinstating the qualitative experience of empathy. These findings need to be further explored and examined empirically.

The trait overlap scores for the AS group were also consistently higher than controls but lower than the DPD group. This pattern was statistically significant for the 'self-overlap' variable and still approaching significance when demographic factors were controlled although this was not the case for 'target-overlap'. These data suggest that the AS group also rely on egocentric processing, as predicted and in line with their 'personal distress' scores from chapter 3. To find that people with AS can use a cognitive representation of the self in this way suggests this early developmental mechanism is intact (see chapter 1), which may be important for devising clinical interventions, if required. However, these results also suggest that allocentric and egocentric simulations are best represented by a continuum rather than discrete categories and for some reason the DPD group fall further along that continuum. This effect needs further exploration as there are clearly differences between the phenomenological experiences of these two groups. Yet if high-functioning people with autism are found to suffer social anxiety (see introduction) there may also be some similarities. It is important to undertake further empirical work to both elucidate the different processes involved in egocentric simulation, and create a model with greater specificity.

Future studies should also focus on the nature of the social cognitive processes that utilise the self-concept. More explanation is needed as to the process that distinguishes egocentric and allocentric simulation. In addition, a more complete understanding of the developmental course of this process may be invaluable for therapeutic interventions. We know that newborn babies can imitate, and the suggestion is that this skill is a precursor to ToM and empathy (Meltzoff and Decety, 2003). We do not know whether simulation or the use of shared representations is an intermediary stage in ordinary development (that remains a useful strategy for understanding affective states – see chapter 4) that may be preserved in milder forms of autism. Furthermore, does egocentric simulation result from participants simply employing executive processes in some kind of compensatory strategy or is it part of the dedicated social cognition module? The association and/or dissociation between both types of simulation and theory-theory processes[21], including their neural and cognitive bases, also needs to be further explored, in order to build a complete picture of the processes involved in social cognition.

4.1 Conclusions

4.1.1 Depersonalisation Disorder and Empathy

The main finding of interest is that people with DPD do not display the same physiological reactions to other people's affective states as other groups. With both sad and happy affective states the reactions they display are in fact opposite to those predicted suggesting that they were not emotionally in tune with the protagonist. This may contribute to their subjective lack of empathy. Their lack of 'personal distress' to others peoples affective states also indicates 'emotional blunting'.

However, the results from the trait task suggest that despite their lack of a physiological egocentric reaction, people with DPD may still have the cognitive underpinnings of personal distress, namely an egocentric representation of others' affective states. The trait task indicates that for the DPD group, intersubjectivity

[21]Although both allocentric simulation and theory-theory processes require a full appreciation of the self/other distinction they differ conceptually as the former also utilises the 'self-concept'.

relies on the 'self' more than it does in control groups. This fits with recent cognitive-behavioural models which suggest that people with DPD have an excessive self-focus.

It is not clear whether these two components of empathy are linked. It may be that facilitating an allocentric viewpoint would automatically lead to appropriate physiological reactions or vice versa but this is not clear from the current data. More studies need to be undertaken to ascertain the relationship between these two processes and to construct a full and comprehensive model of empathy in both healthy and clinical groups. In the meantime, it may be worthwhile monitoring the effect strategies which aim to remove excessive self focus, have on emotional blunting *per se* including affect in response to other people's emotional states.

4.1.2 Asperger's Syndrome and Empathy

The main finding for the AS group is that they can indeed label other people's emotional states and that this seems to be based on a type of simulation rather than purely rule-based reasoning. However, although the evidence suggests people with AS are using an online strategy, it is not possible to derive this definitively from these data alone and replications and further explorations are needed. The data also suggest that people with AS show physiological reactions similar to those of controls (or 'neurotypicals', in the words of PK who has Asperger's Syndrome) and consistent with empathic reactions to other people's affective states, although these reactions seem to stem from the application of a largely egocentric processing strategy.

Although it is exciting to a find this preserved skill in the social cognitive domain, real-life social interaction is not as fractionated as empirical tests, and situations without complex representational states where emotional states appear in isolation, probably occur rarely. People with AS have been shown to do badly on more naturalistic studies of social cognition (see above) and in addition, they may process facial affect in a different way to controls [although see chapter 6 and Blair (2003a)]. Blair (2003a) also points out that ToM is necessary for social referencing which is an integral part of online social interaction. Despite this, it is useful to know that people with AS are relatively intact on this basic social cognitive skill and this may be of help in devising training interventions, if required. This may be especially so for more severe cases if this basic skill (empathy) can be enhanced rather than other more complex methods of attributing mental states.

Dissociating the Components of Empathy

1. Background

A major function of 'theory of mind' (ToM) is the attribution of epistemic representational states, which in some respects require more complex processing than pure affective states. A distinction has been made between the socio-cognitive and socio-perceptual aspects of ToM (Tager-Flusberg and Sullivan, 2000) which rests on the type of stimuli communicating the mental state. The former are said to be specialised for representational mental states and the latter for appreciating mental states from facial and bodily expressions (which are largely, although not solely, affective). This distinction is similar

Contents at a glance

to the earlier division between 'hot' and cold' cognition (see chapter 1) and may have important implications for the type of processes required.

This chapter will describe data that tap both domains, using standardised measures, from people with Asperger's Syndrome (AS) and those with Depersonalisation Disorder (DPD). The AS group were tested on both first and second order socio-cognitive ToM tasks that have been successfully validated on healthy and clinical adult populations (Rowe *et al.*, 2001). In addition, both groups were tested on complex mental state attribution from pictures of the eyes alone – the 'Reading the Mind in Eyes Task' (Baron-Cohen *et al.*, 2001).

1.1 Socio-Cognitive ToM

The phrase 'ToM' was originally used to refer solely to the ability to appreciate representational epistemic states, which are said by some to invoke a theoretical reasoning strategy (see Zahavi and Parnas, 2003). However, it is often used more broadly now (see chapter 1). The idea that people with autistic spectrum disorders (ASDs) have problems appreciating these types of mental states is well established and has been the focus of a large research program (Baron-Cohen, 1999b, Frith, 2001, Frith, 2003) since the 1980's (Premack and Woodruff, 1978, Wimmer and Perner, 1983). More recent work suggests that the appreciation of 'false belief' may be the developmental basis of understanding 'interpretative diversity' or the idea that perception is an active process and 2 people can interpret identical stimuli in different ways (Luckett, 2002). Furthermore, it is now thought that the skills involved in the 'cold' aspects of ToM may well be dissociable from other aspects of empathy, as indicated by the fact that people who display signs of psychopathy tend to do well on such tasks despite deficits in other aspects of empathy (Blair, 1996).

Although the ToM account of ASDs is well documented, it is not altogether clear whether it fully explains the social difficulties shown in ASDs. Although most studies find that people with ASDs do not pass first and second order ToM tasks, some high-functioning individuals with autism pass these tests, including some people with Asperger's Syndrome. However, people with ASDs often reach this developmental milestone at a later chronological and/or verbal age than comparison groups (Baron-Cohen, 1999b, Frith, 2001) and the same people still tend to have problems with more naturalistic (Frith, 2001, Joliffe and Baron-Cohen, 1999, Heavey *et al.*, 2000) or more difficult tasks (Klin, 2000).

One explanation for this inconsistency is the relationship between 'executive dysfunction' and 'ToM'. There is a body of evidence that suggests that people with ASDs may have problems with executive function skills (Hughes *et al.*, 1994). These executive function skills which enable us to 'maintain an appropriate problem-solving set for the attainment of a future goal' (Griffith *et al.*, 1999) and include skills such as cognitive flexibility, planning and set-shifting. The idea is that problems applying these skills underlie the social deficits observed in ASDs. The fact that people with AS sometimes show 'executive dysfunction' but do not always show 'ToM' deficits (Ozonoff *et al.*, 1991) and people with more severe ASDs show problem in both areas, has been taken as evidence for a causal role for executive functions. However, the results have been inconsistent with some young children with autism displaying severe ToM deficits but yet do not display executive dysfunction (Griffith *et al.*, 1999). Furthermore, recent studies have found these two sets of skills to be dissociated in people with selective brain damage (Fine *et al.*, 2001, Bach *et al.*, 2000).

An alternative explanation for the apparently successful performance of high-functioning people with ASDs on ToM tasks, is that they manage to 'hack out' (Luckett, 2002, Happe, 1994a) unusual ways of solving such tasks and employ compensatory strategies (Frith, 2001, Frith and Happe, 1999). This suggests that the apparently preserved skills do not represent an actual sparing of 'ToM'. This is consistent with the observation that people with AS need a higher verbal IQ than control groups to pass such tasks suggesting language may provide 'an alternative neuro-cognitive route to representational thought' (Fisher, 2002). The differential brain activation (as compared to controls) observed in a group of 5 men with AS during first and second order ToM tasks (Happe *et al.*, 1996) sits well with this hypothesis.

Lastly, aside from verbal intelligence, there has been some debate regarding the influence of other aspects of IQ on ToM reasoning (Buitelaar, 1997, Happe, 1994b), although, as already discussed in relation to verbal IQ, it may be a simple case of intelligence masking the deficits in both 'ToM' and 'executive function'. Bearing all these issues in mind, it is surprising that people with Asperger's Syndrome and high-functioning autism are often lumped together in experimental studies, as the diagnostic distinction between the two is based on language *and* intelligence. In fact, Ozonoff (1991) argued that 'collapsing groups with different abilities for the purposes of research projects may obscure relationships and patterns that might be apparent if a more homogenous sample had been used'. If such a 'mixed group' approach is adopted then it is important to match individuals cases to controls using the person-to-person method (Buitelaar, 1997) in order to control for both between group differences and variation within the groups.

In the current chapter, participants with AS were tested on first and second order 'ToM' reasoning tasks in order to further unravel the processes at play in this type of reasoning. These data, when considered along with that from preceding chapters, also allowed the exploration of any association or dissociation between some of the different components of empathy such as ToM and emotional resonance, (Blair, 1996). However, as there was no reason to believe that people with DPD would display problems in this domain ('cold' mental state attribution) they were not given this task.

1.2 Socio-perceptual ToM

People with ASDs may also have particular problems with the socio-perceptual or 'hot' aspects of ToM (see chapter 1 and 5). However, despite evidence of a deficit (Blair, 1996, Hobson *et al.*, 1988) the exact nature is unclear. One study suggests that the observed deficits in facial affect processing may not actually be specific to facial stimuli (Davies *et al.*, 1994). In contrast, another study indicates that the problems may be isolable to social rather than simple basic emotional states (Capps *et al.*, 1992). More recent studies suggest that people with ASDs may rely on different strategies to decode facial affect (Bormann-Kischkel *et al.*, 1995, Hall *et al.*, 2003). For instance, it may be that as a group, people with ASDs pay more attention to the mouth than other regions of the face such as the eyes (Joseph and Tanaka, 2003, Klin *et al.*, 2002). Yet a further alternative view, is that people with ASDs may perform as well as groups that are properly matched for verbal IQ and age (Ozonoff *et al.*, 1990).

The idea that people with ASDs process facial expressions in a distinctive way is partially backed up by recent fMRI studies which suggest they do not activate the traditional 'fusiform face area' when perceiving facial affect (Critchley *et al.*, 2000, Ogai *et al.*, 2003, Pierce *et al.*, 2001). Although one study has implicated the fusiform 'face' area with expertise rather than face processing *per se* (Gauthier *et al.*, 1999). So it is entirely possible that, as a result of their symptomatology, people with ASDs have simply failed to develop expertise in this area. However, the picture is far from clear and in a recent review article, Blair (2003c) argues in line with Ozonoff (1990) – see above - that the findings to date regarding difficulties in emotional processing arise through insufficient matching on mental age. In addition, several studies have not found impairments in the recognition of basic emotion from facial affect – see above plus Adolphs (2001b)[22] – which may mean that the problem is specific to complex emotional states.

Regarding AS, one 'pilot study' (so described because of the incomplete nature of the diagnosis) found 10 people with AS to be impaired on recognising facial affect and affective prosody (Adolphs *et al.*, 2001b). Another pair of case studies included one young man with AS to be impaired on labelling facial affect, but the other performed as well as controls (Shamay-Tsoory *et al.*, 2002). However, one study has specifically addressed the use of compensatory strategies in such tasks and found actual evidence of differential processing of socio-perceptual stimuli, in people with AS (Shamay-Tsoory *et al.*, 2002). Thirteen people with AS and age and verbal IQ matched controls took part in 3 tests of emotion recognition. The first involved the simple matching of emotion words, the second the matching of emotional faces and words, and the last again required the matching of faces and words, however, the faces were already presented with an accompanying word which was sometimes incongruent i.e. happy face with 'happy' and happy face with 'fear'. The people with AS performed significantly worse when identifying the emotion accompanied by an

[22] Although this study did find a deficits in people with ASDs in making 'social judgements' from facial but not lexical stimuli.

incongruent word, despite showing similar performance on the other tasks. This suggests that the verbal information was more salient to them than the subjective affective information. This very much hints at the use of alternative, largely verbal route in solving such tasks

Another study tested those with high-functioning autism and AS on a test requiring mental state attribution from the eyes alone and found a lower accuracy rate than that displayed by controls (Baron-Cohen *et al.*, 2001). It is arguably easier to use unusual compensatory routes to decode facial affect from the whole face and/or basic emotions. The 'Reading the Mind from the Eyes Test', however, tests for more complex mental states with limited cues and so it is less likely that it can be solved using such routes.

Furthermore, using fMRI techniques to measure brain activation during performance on this task, Baron-Cohen *et al.*, (2000) found that a group comprising people with AS and HFA failed to activate some of the same regions as the control group (the amygdala), again indicative of differential processing (Baron-Cohen *et al.*, 2000). On the basis of this and other evidence the authors propose an 'amygdala' theory of autism which finds further backing in the literature (Adolphs *et al.*, 2001b, Howard *et al.*, 2000). Further support for the role of the amygdala in the 'Eyes' task can be found in brain lesion studies (Shaw *et al.*, 2003a). However, it should be pointed out that other areas thought to be involved in more traditional ToM tasks i.e. prefrontal areas were also activated in controls during this task.

The 'Reading the Mind in the Eyes' task has also been given to adults with psychopathy who were not found to perform worse than controls (Richell *et al.*, 2003).

Turning to people with DPD, if we accept that like people with psychopathy, those with DPD may have problems with empathy (see above and previous chapter) and may show blunted physiological responses to emotions, these results may be of help in predicting the performance of the DPD group on this task. However, there is some debate as to how these findings fit with the current literature. There is a body of evidence suggesting that people with psychopathy have amygdala dysfunction (Blair, 2003b) which fits with other accounts of the amygdala as being involved with emotion processing (see chapter 7). If true, and the amygdala is involved in the eyes task, then this would predict that people with psychopathy would find this task particularly difficult, yet this is not the case.

Richell *et al.,* (2003) suggest an alternative explanation for this paradox. They propose that the task recruits the amygdala where possible (and remember that it was only one area in a circuitry thought to be important for such tasks) but that other compensatory cortical routes may be developed to circumnavigate a dysfunctional amygdala if the 'lesion' occurs at an early developmental stage. This explanation is in line with the intact performance of some psychopaths on other 'ToM' tasks (see above), since they are presumed to have developmental difficulties which have been compensated for. The authors also argue that their account explains the severity of deficits observed in children with psychopathic tendencies as compared to psychopathic adults (Richell *et al.*, 2003).

A recent comprehensive study examining a large group of people with amygdala damage occurring at different developmental stages, found that early as opposed to late damage resulted in ToM deficits on socio-cognitive tasks (Shaw *et al.*, 200[illegible]) in press. This is contrary to Blair's assertion that compensatory cortical routes are established in response to early amygdala damage. However, Blair's claim refers to socio-perceptual rather than socio-cognitive ToM and it may be that such compensatory routes are more readily developed to circumnavigate the latter kind of deficit. At this stage the role of the amygdala in different developmental disorders remains unclear, particularly as the theory seems to lack specificity with there being both an amygdala theory of ASDs and psychopathy (see chapter 7 for more thorough review).

In terms of adults with DPD, it is not expected that they will show deficits on this task since despite their subjective problems with empathy these are not generally observed. Even if people with DPD share some characteristics with those with psychopathy, such as emotional blunting and lack of the affective components of empathy (see chapter 5), this would not necessarily interfere with their inability to decode complex mental states. Furthermore, if the amygdala is specifically involved in the perception of complex mental states from the eyes, and it exerts its maximal effect when the damage is early, then there is no reason to believe the DPD group, whose problems have a typically adult/adolescent onset (Shaw *et al.*, 200[illegible] in press.), will be impaired on this task. It is predicted that:

- The AS group will be able to perform accurately first order ToM tasks although they may have problems with second order tasks.
- The AS group will be significantly less accurate than a control group at decoding mental states from photographs of the eyes.
- The DPD group will not show any difference from controls in their ability to comprehend mental states from pictures of the eyes.

2. Methods

2.1 Participants

The same participants will be used as those described in chapter 3 and 5 (see chapter 3, p 2 & 4 for demographic details). The Eyes task was administered to all 3 groups (controls n = 48) and the Rowe tasks to the Asperger's group and all the control participants (n = 53).

2.2 Measures

2.2.1 ToM task

This is a test of advanced 'ToM' ability (Rowe *et al.*, 2001) that has previously been validated on both healthy volunteers and people with frontal lobe damage. It tests both first and second order ToM inference (see appendix 5 and 2) and is aimed at a general adult population. In the original study there were 12 vignettes, 6 first order and 6 second order; however due to time constraints, 4 vignettes were used in the current study - 2 first order and 2 second order. Each vignette was followed by a set of questions:

i) False belief test question - taps the ability to make mental state inferences regarding a false belief.

ii) Inference question – tests the capacity make non-mental state inferences.

iii) Fact question – focuses on facts about the relevant events that are integral to the false understanding.

iv)Memory question – to assess the role of memory in the task.

The data from this task were coded according to the author's original scoring method. The responses for 20 participants were rated independently by 2 people in order to examine inter-rater reliability. For the control sample there was 97.5% agreement for the first order tasks and 92.5% agreement for the second order tasks with 98% agreement for the inference, memory and fact questions. All 16 responses for the clinical group were also coded by 2 independent raters and there was 100% agreement for first order tasks and 97 % for second order tasks.

In order to further explore the data, the control group responses were also coded for 'gist' – 1 point being allocated for an answer being partially correct and 2 points for fully correct. This additional coding strategy was formalised when collecting data from the clinical group, and an additional probe question was included, designed to elicit the extent to which the participant understood the extract

2.2.2 Eyes Test

In the 'Reading the Mind from the Eyes Test' (Baron-Cohen *et al.*, 2001), participants are required to look at a photograph of an actor's eyes and choose one of 4 words that best describes what the actor is thinking or feeling. The task is one of mental state inference from socio-perceptual cues as opposed to verbal material and it has been shown to reliably distinguish between groups of healthy volunteers and those with AS/HFA. Participants were allocated 1 point for a correct answer yielding a total score out of 36.

2.3 Procedure

The first and second order 'ToM' tasks were presented in written format as opposed to orally. Participants were given the extract and told to read it as carefully as they would be required to answer questions afterwards and would not be able to refer back to the extract. The extracts were then presented in a counterbalanced order using a reversed 4 x 4 Latin Square along with 4 diary vignette (see chapter 4) i.e. one vignette followed by one ToM extract.

For all groups, the Eyes task was, however, presented at the end of the testing session. The task was self-paced and participants were required to say their answer out loud as opposed to writing it down. The instructions were adapted slightly to incorporate this change:

For each set of eyes, choose and say out loud the word which best describes what the person in the picture is thinking or feeling. You may feel that more than one word is applicable but please just choose one word, the word which you consider to be most suitable. Before making your choice, make sure that you have read all 4 words. You should try to do the task as quickly as possible. If you really don't know what a word means you can look it up in the definition handout.

3. Results

The dataset for each task was examined for normality and if necessary, the data were transformed using the appropriate method e.g. logarithmic and square root. This was followed by simple analyses to compare group performance on each task. Demographic factors were then explored: gender, education (measured categorically), age, occupation (whether or not the participant was a mental health worker) and verbal IQ were considered. Lastly, a full analysis was conducted to examine group differences on the dependant variable for each task, adjusting for any demographic factors previously shown to be of relevance.

To summarise, for each task the following analyses were conducted:

- A simple analysis to examine group differences on the dependant variable.
- A simple analysis to explore the association between demographic factors and the dependant variable.
- A full factorial analysis to examine group differences on the dependant variable whilst controlling for the necessary demographic factors.

3.1 ToM Task.

Simple group analysis

Initially, mean and standard deviation percentage accuracy scores were calculated for each pair of first and second order ToM tasks in order to allow comparison with the normative data (see Table 6.1).

Table 6.1: Mean and SD accuracy rates on false belief questions

	First order vignettes - no. correct		First order vignettes - % correct		Second order vignettes - no. correct		Second order vignettes - % correct	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CONTROL	1.32	.78	66.04	38.94	.34	.48	17	23.93
AS	.81	.40	40.63	20.16	.38	.62	18.75	30.96

As the data significantly differed from the normative data provided by Rowe *et al.*, (2001) (first order 88%/second order 66%) and this was most marked on the second order tasks, the data were examined more closely. It transpired that only 4 people managed to express the correct attribution on one of the second order extracts. This task was therefore dropped from subsequent analysis.

The remaining data were reconverted to raw scores. They were non-normally distributed so non-parametric tests were used to examine group differences. A Mann-Whitney test revealed significant group differences ($z = -2.27$, $p = .023$) on the test questions and it seems that the AS group found this task more difficult than controls. There were also significant group differences on the non-mental state inference question in the same direction ($z = -1.96$, $p = .05$). However, this was found to be statistically independent of accuracy rate on the test question ($F_{(1, 65)}$ 0.59, $p > .05$) when entered into the analysis as a covariate. There were no group differences on either the memory ($z = -.1.54$, $p > .05$) or fact questions ($z = -1.11$, $p > .05$). Chi squared analyses also revealed no significant group differences on the *type* of attribution made on these 3 tasks.

Demographic factors

Analyses were then conducted to examine the effect of various demographic factors on accuracy rate on the 3 remaining ToM extracts. One instance of missing education data was substituted with the group mean. The percentage of correct answers was not related to age ($t = -0.17$, df 62, $p > .05$), gender ($t = -3.14$, df 62, $p > .05$), verbal IQ – based on a median split at 122 - ($t = 1.55$, df 58, $p > .05$), however, it was related to the presence or absence of postgraduate education[23] ($t = -2.77$, df 64, $p = .07$), and full IQ as estimated from the NART ($t = 2.59$, df 58, $p =.012$).

[23] Presence or absence of postgraduate education showed the largest effect so statistical values are based on this analysis.

Full analysis

Next an ANCOVA was conducted in order to control for the demographic variables. Although parametric assumptions were not met there was fairly large sample size and homogeneity of variance, so a decision was taken to use this test as no equivalent non-parametric method exists. Accuracy rate was the dependant variable and there were 2 between group factors i) diagnostic group (2 levels) and ii) postgraduate education (2 levels) and with full scale IQ as a covariate. There was no interaction between full scale IQ and group ($F_{(1, 60)}$ .83, $p > .05$) suggesting the assumption of homogeneity of regression slope were met and the groups displayed similar relationships between IQ and accuracy rates.

The full factorial model was run and the main effect previously observed for group disappeared ($F_{(1, 60)} = 1.59$, $p > .05$) with the dataset adjusted for full scale IQ (see Table 6.2).

Table 6.2: Between-participants effects on % correct on first and second order ToM tasks adjusted for IQ

Tests of Between-Participants Effects

	Type III Sum of Squares	df	M an Square	F	Sig.
Corrected Model	8985.11	4	2246.277	3.185	.02
Intercept	1083.64	1	1083.643	1.536	.22
IQ	2157.83	1	2157.830	3.059	.086
GROUP	1118.22	1	1118.224	1.585	.213
EDUCATION	8.33	1	8.33	.012	.914
GROUP * EDUCATION	291.99	1	291.99	.414	.523
Error	38792.67	55	705.32		
Total	197777.78				
Corrected Total	47777.78	59			

There was no main effect for education, ($F_{(1, 60)}$ = .012, p > .05) IQ or any interaction (see above). However, full scale IQ as estimated from the NART was approaching significance ($F_{(1, 60)}$ = 3.06, p = .086) suggesting this accounted for the majority of the variance in this task. It seems that despite the fact that these two groups appeared fairly well matched on full scale IQ (AS grp mean = 117.29 ± 5.83 and control mean = 120.56 ± 4.61) there were significant differences on this variable (t = 2.239, df 61, p = .029) which may have influenced performance on this task, hence the earlier significant result.

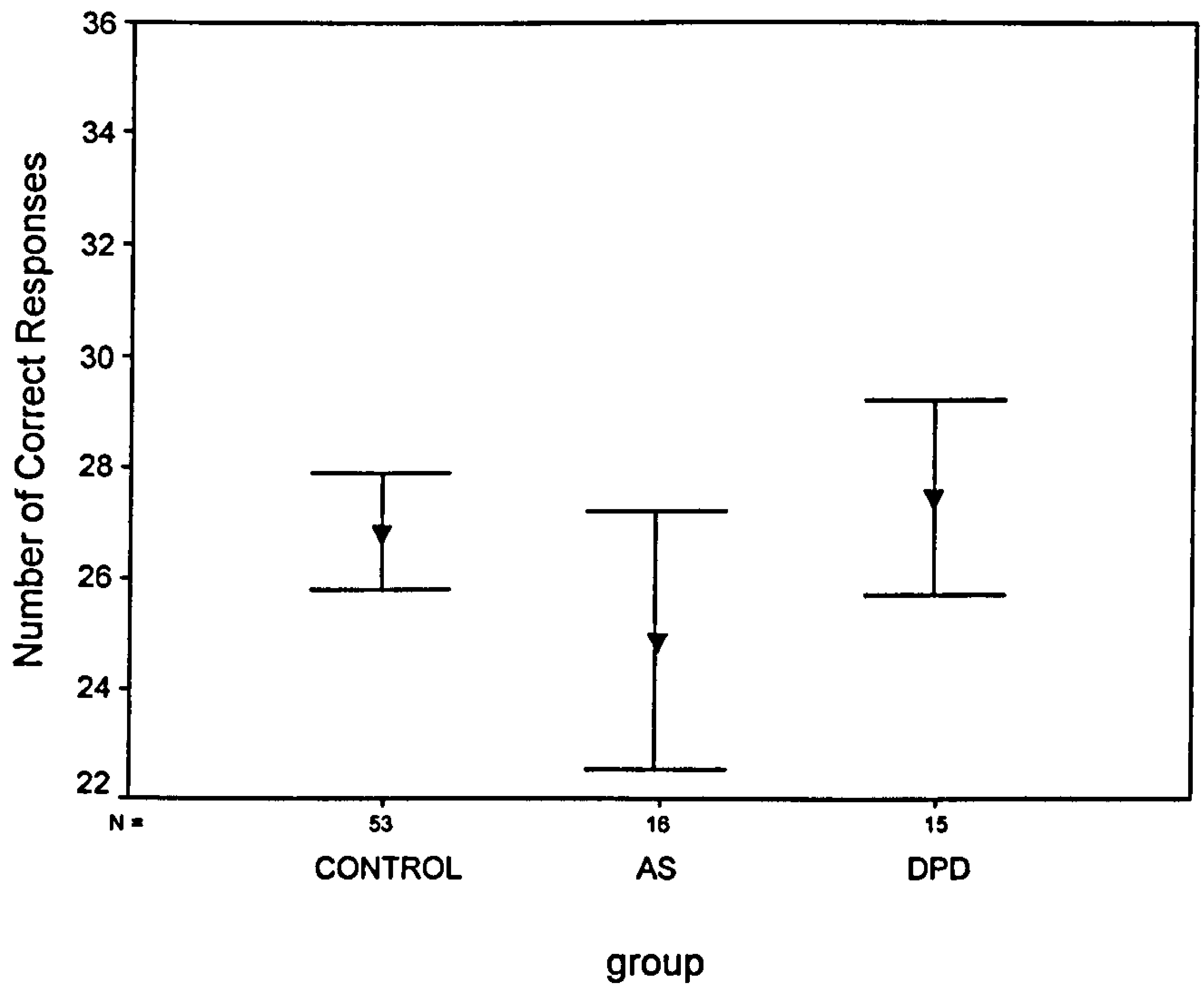
Lastly, the data for the extra coding strategies were transformed to percentage accuracy rate overall for all 4 first and second order tasks and these were found to be 78.1% for the AS group and 70% for the controls.

3.2 The Eyes Test

Simple group analyses

Initially, the total score on the eyes task was calculated for each group (see figure 6.1).

Figure 6.1: Mean and 95% confidence interval on the eyes test



Both clinical groups met the assumption of normality but this was violated in the healthy control group. The dataset was examined for outliers and two cases dropped (one person scoring 16 and another of 19 – see table 3 for means and SD's), however, the sample still remained negatively skewed. But as the sample size was considered sufficient and there was homogeneity of variance, parametric tests were used. A one way ANOVA with *post hoc* Scheffé tests revealed near significant differences between the AS and control groups on the task ($F_{(2, 81)}$ = 2.97, p = .057) suggesting that as predicted the AS group performed significantly less well than controls.

Table 6.3: Mean and SD Score on 'eyes' test

		Mean	SD	n
group	CONTROL	27.22	3.37	51
	AS	24.88	4.4	16
	DPD	27.47	3.16	16

Demographic factors

Next the influence of demographic factors on accuracy rate on this task was examined. Gender (t = -2.25, df 80, p = .027), education (t = -.2.32, df 80, p = .023) and verbal IQ (t = 2.19, df 74, p = .032) all played a significant role in accuracy in decoding mental states from the eyes. To control for the potentially confounding effects of these variables they were all entered in subsequent analysis (the AS group comprised 2 women and 14 men).

Full analysis

An ANCOVA with 3 between group factors (gender: 2 levels, education: 2 levels and group: 3 levels) and 1 covariate (verbal IQ) was performed (see Table 6.4).

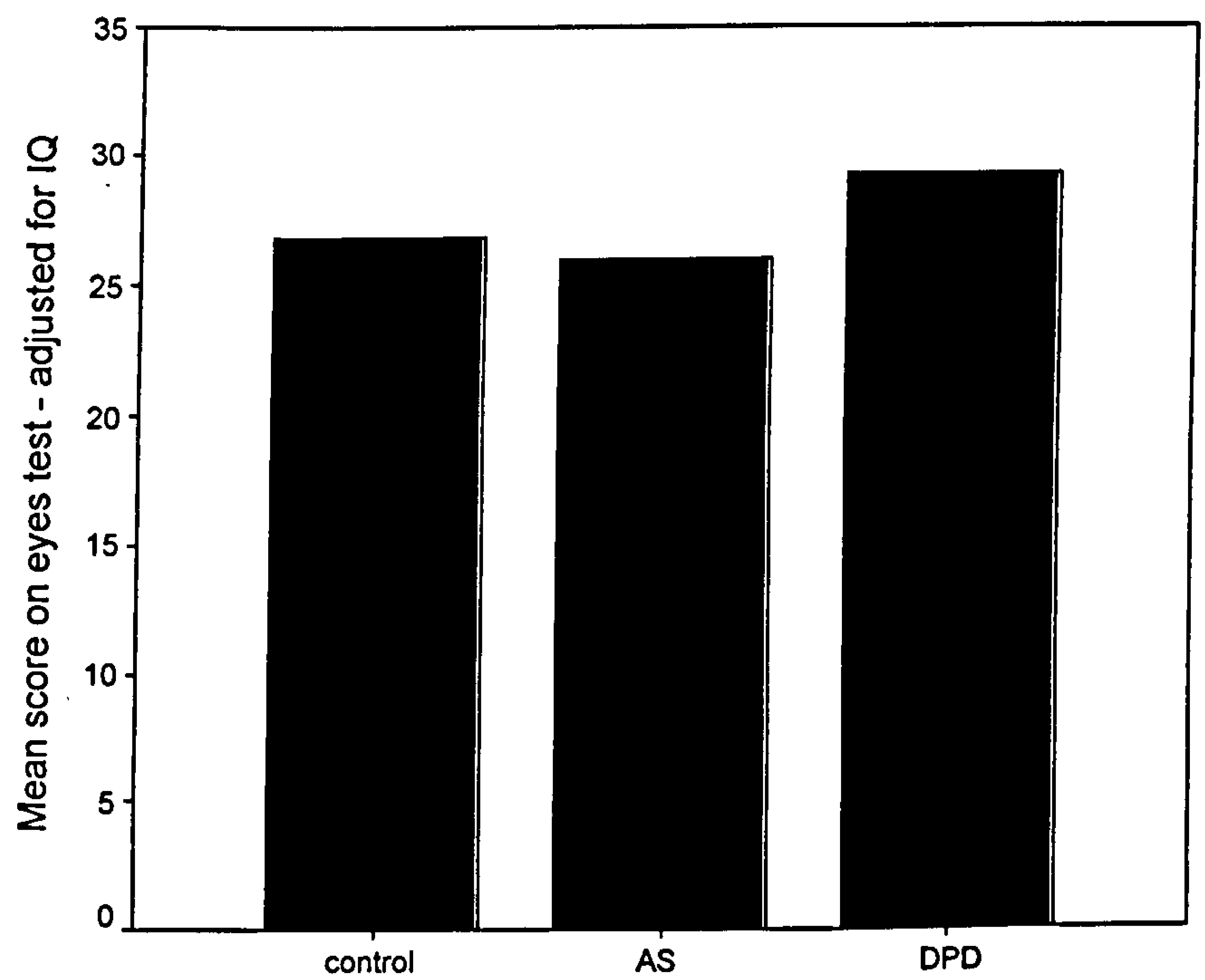
Table 6.4: Between-participants effect on the 'eyes test' adjusted for verbal IQ

Tests of Between-Participants Effects

	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	364.71	11	33.16	3.176	.002
Intercept	4.49	1	4.49	.430	.514
VERBALIQ	135.44	1	135.44	12.975	.001
GROUP	40.13	2	20.06	1.922	.155
SEX	42.6	1	42.6	4.081	.048
EDUCATION	6.74	1	6.74	.646	.425
GROUP * SEX	14.5	2	7.25	.695	.503
GROUP * EDUCATION	.3	2	.15	.014	.986
SEX * EDUCATION	.54	1	.54	.052	.821
GROUP * SEX * EDUCATION	1.43	1	1.43	.137	.712
Error	668.07	64	10.44		
Total	55737	76			
Corrected Total	1032.78	75			

The homogeneity of regression lines assumption was met and there was homogeneity of variance. There was a significant main effect for gender ($F_{(1, 76)}$ = 4.08, p = 0.48) and verbal IQ ($F_{(1, 76)}$ = 12.08, p = .001) accounting for better performance in females and those with higher IQ's. However, group differences were not significant ($F_{(2, 76)}$ = 1.92, p = .155) and neither were any of the interactions (see Figure 6.2). This suggests that the demographic and other general factors were more important in this task rather than group membership. Despite the non-significant effect for group there was a near significant difference using planned contrasts between the DPD and control groups (p = .063).

Figure 6.2 Mean accuracy rate on 'eyes test' adjusted for verbal IQ



4.

Discussion

This group of people with Asperger's Syndrome performed significantly less well than the control group on both the socio-perceptual and socio-cognitive ToM task. However, these effects largely disappeared once the domain general skill of intelligence was fully controlled. The DPD group however performed in line with predictions and did not display any problems on the 'Reading the Mind from the Eyes' task, and in fact they were found to perform particularly well suggesting comparisons between DPD and psychopathy may be fruitful.

The interpretation given for performance on the Rowe *et al,* (2001) 2nd order ToM tasks is hampered by the low accuracy rate observed in healthy controls. For some reason, all participants found one of the second order tasks particularly difficult. Unfortunately, scores on this specific extract were not reported in the normative data, so it is not clear whether or not these results are consistent with the original study. This is compounded by the fact that the original study used 12 vignettes (as opposed to 4), so poor performance on only one extract would be obscured in the grand mean. This disparity and resulting small range of scores may also account for the lower accuracy rates reported overall, even when the particularly difficult extract was dropped.

Alternatively, it is possible that the control group selected were not reasonably representative of the normal population. However, this is unlikely as the groups' performance on other standardised tests (see chapter 3) suggests otherwise. Aside from these possibilities, the only other known difference between this and the original study is that the extracts were originally presented orally, compared to the written format in the current study and that the instructions were adapted to reflect this. It seems improbable that this could account for the discrepancy in accuracy rates.

Despite the lack of consistency regarding accuracy on this test, it is still possible to obtain some insight into the key processes needed to solve such tasks. Group differences were found on both the mental state and non-mental state inference questions. However, none were found once full scale IQ as estimated from the NART was controlled. This lack of group difference once IQ was taken into account, suggests IQ plays an important role in 'ToM' tasks and that the previously observed group differences are likely to be due to domain general rather than specific processes. Although the mean IQ for the AS group was quite high in this study, there were significant group differences on this variable and approximately 20% more variance in the clinical group. It is therefore possible that the near significant effect for IQ partially resulted from the wider range of scores in the clinical group. If so then a reasonable explanation for these findings is that those with high IQs in the clinical group were able to apply compensatory strategies (see above) in line with the 'hacking out' hypothesis. This option may not be available to people with sub-acute focal brain damage.

Interestingly, these findings are not necessarily inconsistent with Rowe and colleagues' original study in which the authors found no significant group differences on verbal or full scale IQ. In that study, it appears there may have also been more variation in IQ in the clinical groups (control SD 8.3, cases SD 13.35). This suggests that individuals were not specifically matched on this factor and it is therefore possible that this failure to match cases and controls pairwise may have accounted for some of the variance in error rate. Furthermore, IQ was not entered as a covariate in the final analyses.

In addition, the discrepancy between the original and additional coding strategies employed suggests that *both* groups understood more than they managed to express – or failed to express their knowledge in explicit 'mentalising' terms as required by the experimenter i.e. 'he thinks she thinks' etc. It also seems that the scoring procedure on this task was perhaps more stringent than other similar tasks (Happe, 1994a) as items were automatically considered incorrect if they did not relate very closely to the prototypical answer. Unfortunately, tasks such as these lose the objective and definitive qualities of first order false belief tasks, as accuracy rates are to some extent based on the ability of participants to express their 'mentalising' ability in a particular and specific way. In support of this idea, some argue that the use of such language is indicative of 'spontaneous' mentalising (Happe, 1994a), however, it is also possible that such tasks tap other skills such as the pragmatic language ability with which people with ASDs and frontal lobe damage (Rowe *et al.*, 2001) may have problems. In one study requiring the attribution of mental states to geometric shapes - see above - (Klin, 2000) people with AS (as opposed to other ASDs) were found to improve their performance when additional, more explicit, instructions were given.

In the current paradigm, however, the authors in the original study ruled out the latter explanation because although the clinical group answered the non-mental state inference questions incorrectly, which they took to be a measure of pragmatic language skill, performance on these items was statistically independent from the main effects. The same pattern of results was also found in the current study. However, this line of reasoning is based on the assumption that pragmatic language ability is equal for both sets of questions i.e. both mental and non-mental state inference questions and it is not entirely clear that this is the case. It is in fact possible (and seems likely) that the non-mental state inference questions are less ambiguous, with their pragmatic purpose being more explicit rather than in the test questions, for example:

Test Question: Why does Richard say this?

Control Participants Responses: 'Because he assumes she hasn't got the message, and so thinks he hasn't been home and started'.

Inference Question: Why do you think Ann needed to hurry to get to the DIY shop?

Control Participants Responses: 'Because it closed at 6pm.'

If this is the case, and the mental state items tap pragmatic language ability more so than the non-mental state items, then such language ability would be a potential confound. In the current study, the group differences were larger for the mental state inference items which by this line of reasoning could reflect the role of increased ambiguity and the need for pragmatic language ability. This explanation is consistent with the findings from the alternative coding system which yielded higher accuracy rates, and the fact that the control group showed a lower accuracy rates than expected, suggesting they also found the items ambiguous.

Interestingly, the opposite of failing to appreciate the pragmatic cues of a situation, is a well known phenomenon within psychology known as 'demand characteristics'. This is the ability of more socially skilled people to pick up the requirements of the experimenter that have only been implied, although this is not restricted to language. It is possible that this may have occurred with some individuals in the control group. In replication, more explicit instructions and unambiguous questions, perhaps with examples, may lead to a higher accuracy rate across groups. Furthermore, due to the subjective nature of such tasks it is important that co-validation be performed with at least one rater blind to the clinical group which, unfortunately, was not possible in the current study.

Turning to the 'Reading the Mind in the Eyes' task, the finding that people with Asperger's Syndrome had a lower accuracy rate than controls on this task is consistent with previous findings. Again this effect disappeared once IQ was statistically controlled. This suggests the task again taps domain general rather than specific skills. In the original study, verbal IQ was not controlled in all the comparison groups, but a comparison group matched for verbal IQ was selected, and group differences were still found. Again it is not clear whether the individuals were matched for IQ or the group overall, although the SD was larger in the clinical group, suggesting more variation.

A further difference between this and the original study is that the latter did not statistically control for gender differences. If verbal IQ and gender differences do play significant roles in performance on the eyes task, then perhaps these factors account for some of the variance in the original study as opposed to simply domain specific group differences. There are two possible explanations for the role of verbal IQ in this task, firstly, that it is the basis of compensatory strategies in the ASD group (as with the socio-cognitive task) or secondly, that it exerts an influence through the verbal load of the task. The latter is unlikely as participants were provided with a handout detailing definitions and examples of all the words used. It seems more probable that this group of people with AS, who had particularly high IQ's, were also able to put that intelligence to use in solving the task using unusual strategies. In replication, it may be worth coding the data according to the more sensitive coding strategy suggested by Adolphs *et al.*, (2002). This approach uses the actual responses of the control group as a baseline rather than the data gathered when the task was developed

The fact that IQ accounted for much of the variance in both 'ToM' tasks does suggest that people with AS and high IQs can construct and employ compensatory strategies to meet the requirements of such experimental tasks. This is in line with previous studies and highlight the gap between performance on empirical tests that are purposely fractionated for experimental control and real-life online situations with rapid and dynamic stimuli (Klin *et al.*, 2000).

The DPD group performed as well as controls which is in line with the performance of people with psychopathy, suggesting this group can decode complex affective states. Furthermore, despite a non significant main effect there is some evidence to suggest superior performance once IQ and gender were controlled. To find that this group may be particularly good at this task, whilst failing to show physiological arousal (see chapter 5) suggests a profile that is similar to psychopathy. Although, there are obvious phenomenological differences between the two groups, comparisons may be useful theoretically.

Lastly, if the data from the AS group are taken to represent an instance of successful compensatory strategies rather than spared 'ToM' ability, then the possibility of dissociation between some aspects of 'affective empathy' becomes tenable (Blair, 1996). In chapter 5, the suggestion was that those with DPD did not have the emotional arousal necessary for 'affective empathy' in contrast to the AS group. However, in this chapter, the AS group seem to use unusual processing routes to decode complex mental states whereas the DPD group displayed no problems with this aspect of empathy. This may provide additional evidence of a 'double dissociation' (Blair, 1996) between some components of empathy i.e. 'ToM' and physiological response, and is in line with earlier evidence suggesting a dissociation between psychopathy and ASDs (Blair, 1996). This does not, however, stretch to the other aspects of empathy identified such as labelling affective states from non-facial stimuli and the use of the self-concept.

In replication, it would be useful to employ more sensitive indices of the strategies employed such as reaction times and fully match individuals within the groups on IQ, gender and education. This is especially true since the control group in the current study may not have been representative of the normal population regarding education, although there was no significant main effect for education in any of the analyses. It may also be helpful to formally gather qualitative and subjective data as to the nature of the processes employed, especially as the groups detailed here probably have sufficient insight. Regarding the socio-cognitive ToM task more explicit task instructions as to the nature of response required to get the question correct, with perhaps a few examples beforehand, may prevent responses being scored as incorrect simply on the basis of ambiguity. This would have to be balanced against the risk of inducing a response set as a result of priming the participants.

4.1 Conclusions

In summary, it appears that people with Asperger's Syndrome may employ unusual strategies to pass first and second order ToM reasoning tasks in an experimental setting. However, it is not clear how much these tasks tap additional skills such as pragmatic language ability. People with AS were also found to perform at the same level of controls on the 'Reading the Mind in the Eyes' task once factors such as gender and IQ were controlled. Again, due to the variation in IQ in the clinical group, including a number of individuals with quite high IQ's, it is likely that this is due to reliance on complex and unusual rule-based compensatory strategies.

To find that people who report DPD did well on the 'Reading the Mind from the Eyes' task suggests that their emotion perception skills are intact despite their lack of physiological responses to other people's affective states. This provides further support for validity of comparisons between psychopathy and DPD. Lastly, data presented in this and the last chapter may go someway to suggest a 'double dissociation' between these two components of empathy (Blair, 1996). The DPD group seem to have preserved 'ToM' ability but may lack the affective physiological component of empathy, whereas the AS group have may have the affective component intact but the 'ToM' component is impaired.

The Neural Correlates of Empathy

1. Background

Shallice (1989) has argued that multi-level theories in psychology have more predictive power than single level theories. In other words, the more we know about a process from different levels of explanation, the more likely we are to have an accurate understanding of use in explaining both existing and novel conditions. As well as adding to existing knowledge, localisation of cognitive functions to brain regions can also play an important role in 'falsification' (Popper, 1963) and in helping to decide between competing theories (Shallice, 2003). Establishing the neural correlates of empathy and investigating how the different components may be fractionated, will therefore be of value in establishing a comprehensive and

Contents at a glance

clinically applicable model. Cognitive functions can be localised within the brain through the study of people with selective brain damage – a method that has been used successfully for many years. However, relatively new techniques allow the mapping of cerebral blood flow and hence function, *in vivo*. Functional magnetic resonance brain imaging (fMRI) is particularly non-invasive as it relies on the magnetic properties of oxy- and deoxyhaemoglobin as opposed to the radioactive tracers used in other methods i.e. positron emission tomography (PET). *In vivo* techniques again play an important part in the 'falsification' process by providing new evidence for theories previously reliant on data from 'lesion' studies alone. In addition, these methods can tell us about the *healthy* as opposed to the injured non-typical brain, which may differ in important ways (Savoy, 2001). Such techniques may also be especially helpful in situations where processes are difficult to detect using simple cognitive paradigms i.e. when people use compensatory strategies (see chapters 5 and 6).

This chapter describes data gathered from a newly adapted test of social cognition designed to test the neural correlates of mental state attribution using naturalistic and hence more ecologically valid stimuli. The test is based on a well-validated and dynamic test of non-verbal sensitivity (Rosenthal *et al.*, 1979). For the current study, it was computerised, updated, and revised to suit an fMRI paradigm.

1.1 Socio-cognitive ToM

There is now a large body of research looking at the functional neuroanatomy of socio-cognitive ToM (Tager-Flusberg and Sullivan, 2000) or 'cold cognition' (Stone *et al.*, 1998), that is, the ability to attribute representational mental states (see Frith and Frith, 2003 for a review) without reference to feelings and emotions. One of the first studies was conducted by Fletcher and colleagues using a PET paradigm during which participants interpreted non-literal utterances, in the Happé Strange stories task (Happe, 1994a). Activation specific to the experimental condition was found in the left medial frontal gyrus and posterior cingulate cortex. A series of further studies using the same task also found the medial prefrontal areas to be activated (Vogeley *et al.*, 2001, Gallagher *et al.*, 2000). In addition, Gallagher *et al.*, (2000) used similar stimuli in cartoon form designed to elicit representational ToM which again activated medial prefrontal areas along with the superior temporal sulcus and temporal poles. Frith and Frith (2003) point out that these 3 areas are consistently activated during the performance of such tasks.

The medial prefrontal cortex has also been shown to be activated in studies that induce implicit mentalising such as a game of stone, paper, scissors (Gallagher *et al.*, 2002) and in a game involving mutual co-operation (McCabe *et al.*, 2001). Of interest is that this activation was specific to the condition where participants thought their opponent was a person rather than a computer, suggesting these areas are involved in detecting intentionality[24]. Further tasks that evoke neural activation in medial frontal areas include a study using the Heider and Simmel (1944) paradigm in which participants view geometric shapes moving on a screen and then attribute mental states to them (Castelli *et al.*, 2002). Lastly, the medial prefrontal cortex has also been implicated in imitation (Decety *et al.*, 2002) which is consistent with recent ideas that imitation my be a precursor to empathy (Meltzoff and Gopnik, 1993b).

[24] Baron-Cohen proposes the existence of an intentionality detector which attributes intentions to agents that are moving and/or acting on objects.

A further study purporting to examine the neural correlates of empathy has also used stimuli that tap socio-cognitive ToM. Farrow *et al.,* (2001) had participants make a series of social judgements including some requiring empathy (understanding and appreciating the protagonists' emotional state) and forgiveability (deciding which crimes are more forgivable). Main activation for the empathy condition was found in the left anterior middle temporal gyrus.

Furthermore, Eslinger (1998) review data from their earlier study where a group of 40 people who had suffered damage to their frontal lobes, undertook a series of empathy self-report measures (Grattan, 1994). As a group, they were found to score lower than the comparison group on items tapping both cognitive and emotional empathy (see chapter 2) although the authors conclude that this may be for different reasons. Inverse associations were found between measures of 'cognitive flexibility' and empathy suggesting that this executive function is key. These correlations were most evident for the people with left dorsolateral lesions. This led the authors to tentatively suggest that the dorsolateral prefrontal cortex may be specifically involved in perspective-taking or cognitive empathy by virtue of cognitive flexibility whilst the orbitofrontal cortex be more specialised for emotional empathy (also see Shamay *et al.*, 2003). This is also consistent with the somatic marker hypothesis (Damasio, 1994).

In addition, Stone *et al.,* (1998) found that people with bilateral orbitofrontal lesions had problems on a test of the *faux pas* whereas those with unilateral dorsolateral prefrontal cortex (DLPFC) did not. Detecting a *faux pas* requires the ability to integrate mental state attribution with empathy i.e. appreciating that the actions of the protagonist could have been hurtful, and it seems this may be where the problem lie. It seems therefore that orbitofrontal areas may play a role in this integrative aspect of empathy, fusing an appreciation of the epistemic and affective mental states of others.

1.2 Socio-perceptual ToM

The neural correlates of socio-perceptual ToM or 'hot cognition', and in particular regarding facial affect, have been extensively characterised. One set of theories argue that the affective processing of each of the basic emotions is subserved by a dedicated unique neural substrate. Blair & Perschardt (2003) in response to Preston and deWaal (2003) point out that this possible fractionation of emotion perception has important implications for localising the neural correlates of empathy, assuming that empathy entails, in part invoking a similar mental state to that observed – be it fear, disgust, joy etc. The evidence for this position is most compelling for the role of the human amygdala in the processing of frightening stimuli. Several studies of people with bilateral lesions of the amygdala have found deficits in the recognition and evaluation of fearful faces (for instance see Adolphs *et al.*, 1994). This is complemented by fMRI studies which have demonstrated activation of the amygdala, particularly in paradigms which involve the passive viewing of fearful faces (for instance see Whalen, 1998).

But whether the amygdala is specialised simply for fear, all negative emotions or those which evoke withdrawal behaviours is unclear. Examining the lesion studies there is considerable variation in impairments in the recognition of facial expressions of emotion arising from bilateral lesions of the amygdala, with some studies reporting deficits in all emotions, and others reporting no impairments at all relative to brain damaged controls (Hamann and Adolphs, 1999). Similarly, amygdala activation has been reported to not only fearful faces, but also facial expressions of anger, sadness and even happiness (for a review see Zald, 2003).

Further evidence in favour of the categorical perception of emotions, comes from the impairment of disgust processing observed in people with pathology of the basal ganglia (such as Huntington's disease) and insula. The insula has been found to be selectively activated in response to facial displays of disgust (Phillips *et al.*, 1998). In addition, one recent study found the same areas to be activated in the same participants when experiencing disgust (from inhaling a noxious odour) and perceiving disgust (from seeing a dynamic disgusted face) – see Wicker *et al.*, (2003a). However, the insula has been found to be responsive to emotional states *per se* and recent theories propose this area may be useful in reconstructing an online representation of the perceived emotional state (Adolphs *et al.*, 2000) and it may be that this process is particularly necessary when perceiving disgust.

Brain structures may also be specialised to respond to different types of emotional cue. Hariri (2002a) found the amygdala to be more specialised for facial expressions rather than emotionally-laden scenarios. Furthermore, in a pair of experiments, Hoffman and Haxby (2000) found that face identity and eye gaze processing had different neural substrates. Eye gaze was found to be mediated by the superior temporal sulcus (STS) with the interparietal sulcus encoding the direction of that gaze. A further study found the STS to be involved in gaze shifts towards the participant (Pelphrey *et al.*, 2003). The superior temporal gyrus has also been implicated in emotion perception under conditions of direct gaze (Wicker *et al.*, 2003b). Others have also found the amygdala to be more activated when reading mental states from direct rather than averted gaze (George *et al.*, 2001, Emery, 2000, Kawashima *et al.*,

1999) which fits with the idea of its role in scanning the environment for stimuli of significance.

Despite the abundance of work on perception of emotion from the facial area, there is a lack of work on the perception of affect from bodily expressions. Sprengelmeyer (1999) used a test of emotional postures with a person with bilateral amygdala damage and a left thalamic lesion. The task required the participant to identify 6 basic emotions (including surprise) from 60 static photographs minus faces of actors portraying the different affective states (10 per emotion). The participant was significantly impaired in recognising fear (as compared to 10 healthy controls). However, Adolphs and Tranel (2003) found the amygdala to only be involved in stimuli that include faces. In this study 4 people with bilateral amygdala damage were found not to be impaired in recognising anger from stimuli excluding but were impaired when the face was present. Furthermore, the people with amygdala damage did not show the usual advantage in recognising emotions (including fear) with the face present.

Aside from these studies, as far as we are aware there is only one study that specifically examines the neural correlates of the perception of mental states from static bodily expressions and gestures alone. Hadjikhani and DeGelder (2003) looked at static fearful bodily expressions and found activation in the fusiform gyrus and amygdala. In addition, there is a growing literature on action perception from dynamic stimuli which is likely to be of relevance (Decety, 2002). That emotion perception is ordinarily based on dynamic stimuli is another important characteristic of real-life mental state attribution (LaBar *et al.*, 2003). Adolphs and Tranel (2003) point out that the use of dynamic stimuli will probably engage different processing streams such as the extrastriate visual cortices and may rely less on the amygdala. However, the use of dynamic stimuli depicting mental state information in localisation studies is in its infancy. LaBar *et al.*, (2003) found activation specific to the superior temporal sulcus when comparing dynamic facial emotion stimuli with dynamic facial identity stimuli. This is consistent with the role of the superior temporal cortex in perceiving 'biological motion' (Adolphs, 2003).

Another study systematically examined this issue and indeed found different neural codes for the perception of static vs. dynamic facial expressions (Kilts *et al.*, 2003). Furthermore, Adolphs *et al.*, (2003) in a study designed to look at the fractionation of emotion systems by valence (see above) found that a person with extensive brain damage that largely spared the somatosensory cortex could not recognise any basic emotions except happiness from static displays but could identify the same emotions from dynamic displays. This again suggests that dynamic stimuli recruit different neural processes to static stimuli and the authors hypothesise that damage to the anterior temporal lobe and medial frontal cortices may be responsible for the deficits observed in this person. Thus, these areas are responsible for associating static stimuli with the brain structures ordinarily used for emotion perception. It seems that dynamic stimuli may invoke different strategies for decoding mental states.

Support for this assertion comes from one study that used dynamic stimuli to examine the neural correlates of 'sympathy' which by our definition would be called empathy (Decety and Chaminade, 2003). Participants were required to watch a series of video clips showing people telling sad and neutral stories that they had supposedly experienced. The experimental manipulation involved the use of stories which were either congruent or incongruent to the emotional expression displayed. In addition to activation in emotion processing areas, there was activation consistent with a 'shared representation' network (Decety and Chaminade, 2003). The idea of 'shared representations' is closely allied to the growing body of work on 'action perception'.

1.3 'Shared representations'

The discovery of neurones in the premotor cortex that are activated in response to both observing and executing an action may be important for developing a model of mental state perception (Gallese and Goldman, 1998). According to Gallese (2003) studies examining action execution and action observation in humans suggest that these neurones may extend back to the inferior parietal cortex which has connections to the superior temporal sulcus. There is also some evidence to suggest that these neurones also exist in Broca's area (Heiser *et al.*, 2003 ~~(in press),~~ Hamzei *et al.*, 2003), an area traditionally thought to be specialised for language. This is consistent with the idea that language evolved from gestures. Gallese builds on the existence of such neurones as evidence for his 'shared manifold hypothesis' which at the phenomenological level is described as 'how our actions, emotions and sensations become implicitly meaningful' and at a functional level relies on the use of 'as-if' modes of interaction to enable models of self and other.

A similar concept has also been proposed based on the use of 'shared representations' wherein 'perception and action share common neural and cognitive codes' (Decety and Chaminade, 2003, Decety and Sommerville, 2003). Grezes and Decety (2001) undertook a meta-analysis of approximately 30 studies examining action generation, action simulation, action observation and verb generation. Many areas were found to be involved but the main areas of activation specific to action simulation studies were the premotor cortex in 83% of studies, and the dorsal part of the dorsolateral prefrontal gyrus, whereas the supramarginal gyrus was activated in all the studies involving simulation, and the superior parietal lobe in 50% of studies. This lends support to the idea that action perception, observation and simulation share common neural codes is therefore well supported.

One ingenious study exploited the 'sharedness' of self and other representations and had participants move their arm whilst observing incongruent arm movement. Interference was found only when the other arm was human as opposed to robotic (Kilner *et al.*, 2003). In another study exploring action, participants were required to simulate actions with either a first or third person perspective. Results indicate that the right inferior parietal, precuneus and posterior cingulate be involved in third person perspective when compared to first person perspective, and the somatosensory cortex and left inferior parietal cortex in first person as compared to third person perspectives (Ruby and Decety, 2001). Activation in the supplementary motor area (SMA), precentral gyrus and precuneus were common to both conditions. However, as the precuneus was more activated in the third person perspective the authors go onto argue that this area be 'critically involved in discriminating self from others, by way of involvement in the representation of the self'. In addition, a further study found the somatosensory cortex to be involved in distinguishing stimuli generated by the self or another, finding less activation in this area when the stimuli was self-produced (Blakemore *et al.*, 1998a).

A later study by these authors found the inferior parietal cortex to be involved both in imitation and being imitated (Decety *et al.*, 2002). Chaminade *et al.*, (2002) also found this area to be activated during an imitation paradigm along with the superior parietal lobe. It seems that the parietal lobe may be key to distinguishing actions and behaviours that belong to the self and others (Blakemore, 2003b). Activation was also found in the dorsolateral prefrontal cortex (DLPFC) during imitation which the authors suggest is due to involvement in the 'preparation of forthcoming action based on stored information'. The parietal areas were also found to be more activated when participants were hypnotised to attribute their own actions as external in a PET study (Blakemore *et al.*, 2003).

Although the notion of 'shared representations' has recently been extended to facial actions portraying mental states (see Decety and Chaminade, 2003 above) there is less work examining whether this model is appropriate for mental state attribution and whether actions depicting mental states i.e. facial expressions and gestures, are processed in the same way as other actions The superior parietal networks bilaterally were activated in a study examining intentional contingency between shapes at a distance whereby intentionality was experimentally induced by having one shape appearing as if it was looking at or following the other (Blakemore, 2003a). Furthermore, Blakemore and Decety (2001b) point out that the fact that the brain appears to be hardwired to perceive 'biological motion' and that we tend to attribute intentionality to such stimuli, indicates a link between the perception of action and the attribution of mental states. Further evidence can also be found in the developmental literature and the observation that infants imitate the expressions of others from a very early age (Meltzoff and Decety, 2003).

Carr and colleagues (2003) conducted a similar study explicitly examining both imitation and observation. Participants were required to either imitate or simply observe a series of photographs (hence static) of facial affect. Both conditions were found to activate a largely similar network in the brain, however, the premotor areas including inferior frontal cortex, superior temporal cortex, insula and amygdala were all activated to a larger degree in the imitation condition. The authors argue that the insula plays an important role in this type of empathy. In addition, one recent study showed cognitive conceptual perspective-taking to activate action perception and recognition areas (Ruby and Decety, 2003) suggesting this strategy may also be useful in general perspective-taking.

The Perception Action Model proposed by Preston and DeWaal (2003) is a broad account incorporating the idea that social cognition stems from the tendency to activate one's own representations in response to perceiving a particular state in another (see chapter 1). This model provides a framework which incorporates much of the work described above. This account finds support in a PET study examining the neural correlates of imagining your own and someone else's emotional experience (Preston, 2002) whereby very similar brain areas were found to be activated. In addition, a further more recent model of emotion perception, draws on early theories of the nature of emotion, and in doing so emphasises the production of an affective state in response to the identification of stimuli of emotional significance (Phillips *et al.*, 2003). Again, this model sits well with simulation accounts.

Further neuropsychological evidence in favour of an action perception or simulation model of mental state attribution comes from another research group who place great importance on the contribution of somatosensory areas in decoding affective states (including the insula). They found data consistent with the use of 'shared representations' including the recruitment of the insula and the somatosensory cortices including supramarginal gyrus (Adolphs *et al.*, 2000) in emotion perception. The researchers equate these findings with evidence for a simulation account (Adolphs, 2003, Adolphs, 1999).

The most comprehensive study to suggest the existence of such a mechanism involved 3D lesion mapping in 108 people with focal brain lesions (Adolphs *et al.*, 2000). Lesions were registered in common brain space to allow comparison and participants were tested on a series of tasks. Damage to somatosensory related areas were found affect naming emotions and appreciating emotional concepts (tested by having participants sort photographs of facial affect according to similarity). The insula alone was specifically implicated in conceptual knowledge independent of lexical stimuli. The authors interpret these data as consistent with the idea that covert somatosensory online representations are constructed when emotion judgments are made.

Adolphs (2002) points out that such a simulation-type mechanism (see chapter 1 and 4) may be particularly useful when the stimuli to be decoded are difficult or when the 'explicit knowledge triggered...is insufficient to recognise the emotion' (Adolphs, 2002). Under these circumstances it might be helpful to construct knowledge regarding the observed mental state via the generation of a simulation. He further argues that facial expressions can be viewed as a special class of action i.e. 'facial affect action' and so it is reasonable to assume that the same mechanisms used in action are recruited in decoding facial affect. However, he also suggests that it is unlikely that this is the case in all instances of facial affect recognition and that this might be dependant on the demands of the task. He describes a useful analogy between retrieving knowledge regarding the shape of your house which may not require the visual cortices at all whereas retrieving knowledge of the number of windows or shape of the windows would require more activation of the visual cortices.

A paradigm that uses particularly difficult stimuli (i.e. fleeting, impoverished or non-prototypical) may therefore be of use in examining the cognitive and neural correlates of 'simulation'. In real life mindreading, social cues are delivered in varied conditions, and sometimes there are lots of rich cues like in a one-to-one conversation but at other times the information available is very limited, for instance, when trying to read the mental state of someone across a room or when a person turns their head away.

As discussed in previous chapters (see chapter 4), the main cognitive difference between all these accounts which can loosely be grouped together as 'simulation' theories and those based on applying a theory, is the use of information about the self in appreciating other peoples' mental states in the former. To this end the neural correlates of self-reflection become relevant. Vogeley *et al.*, (2001) used the Happé strange stories (see above) but manipulated whether or not the participants were protagonists in the story. Based on this and other data Vogeley and Fink (2003) argue that in addition to medial prefrontal areas, medial parietal and lateral temporoparietal cortices are involved in the first person perspective and adopting an egocentric frame of reference. Johnson *et al.* (2002) also found posterior cingulate activation when participants responded to a series of statements about their own abilities and traits. Self-referential processing of trait adjectives also activated this area (Kelley *et al.*, 2002). In addition, Kircher and colleagues (2000) used pictures of the self and partners faces' and found activation in the left inferior parietal areas when viewing the self. Moreover, Kircher *et al.*, (2002) found self-referential processing to activate the left superior parietal lobe and left precuneus. These data are consistent with the idea that parietal areas are important for distinguishing self from other. However, although these brain areas are likely to be important, a recent study by Ramnani and Miall (2004) provide evidence to suggest that the overlap between self and other isn't complete. They found differential activation in the premotor cortex in response to adopting a third and first person perspective and argue for a mechanism based, for instance, on mental imagery of the other's actions.

To conclude, it seems that several unresolved issues remain that are of relevance to establishing the neural correlates of empathy. Firstly, there is a lack of studies using dynamic stimuli to portray mental states and hence exploring whether an action perception model is appropriate for decoding mental state 'actions'. In addition, it is not known whether such a model is only applicable to stimuli that are particularly difficult to decode or with all dynamic mental state expressions. Furthermore, it is not clear to what extent areas relating to self representation are recruited in decoding others' mental states. Lastly, data are scarce as to the neural correlates of mental state attribution from bodily expressions alone.

The Profile of Nonverbal Sensitivity - PONS - (Rosenthal *et al.*, 1979) is an offline mental state attribution task based on video clips of an actor displaying various mental states using facial and bodily expressions and gestures. It is therefore well placed for examining these issues. The PONS is a well validated test of a person's ability to decode mental state information without verbal cues (see appendix 7 for description of stimuli). It uses a series of 2 second video-clips of an actor in various affective interpersonal situations and hence displaying expressions and gestures depicting different mental states. These clips are presented in 3 different physical channels: i) face ii) body iii) figure (face and body together – see appendix 8 for examples). The task is forced choice from 2 options, and the similarity of the alternative is systematically varied, hence introducing an additional 'difficulty' variable. This task has been adapted for an fMRI paradigm by adding 2 control conditions and computerising it (Lawrence *et al.*, 2004). The mental state labels have also been refined and the stimulus presentation counterbalanced more thoroughly. Over the years the PONS has been used with many different samples and there are a wealth of data confirming its reliability and validity in both healthy individuals and clinical groups. Furthermore, it has recently been revived and used in a study with people with schizophrenia who were shown to perform less well than controls (Toomey *et al.*, 2002). This revised version of the online (online PONS-r) has been piloted 'offline' on 20 healthy volunteers. It is predicted that:

- The PONS-r will activate brain areas associated with 'shared representations' such as premotor and SMA, inferior frontal gyrus, precentral gyrus, somatosensory areas such as the insula and supramarginal gyrus, and areas in the parietal cortex such as inferior parietal cortex, precuneus, and superior parietal lobule.
- The medial frontal cortex may also be activated during the PONS-r in line with traditional ToM tasks.
- Areas relating to self processing such as the precuneus and posterior cingulate will also be activated during the PONS-r.
- The amygdala will not be activated when participants view the clips in the PONS-r as at least 50% of the clips do not include the face and in the clips that do, eye gaze is not directed at the participant.
- If a 'shared representations' network is only activated with dynamic stimuli that are difficult to decode then the 'figure' stimuli will not activate this network.
- The fusiform gyrus will be activated in response to 'body' stimuli.

2. Task Development

2.1 Stimuli

The original version of the Profile of Non-Verbal Sensitivity (Rosenthal *et al.*, 1979) contained 20 clips of real-life situations presented in 11 conditions along with instructions to 'judge which of the two real-life situations is represented by the segment'. There were 2 sound conditions one with the content distorted and the other with the tape randomly spliced to obscure the content. There were 3 visual non-mental state inferences, face only, figure and body only. Each non-mental state inference was presented without sound (3 conditions) and with each of the two sound conditions (6 conditions) and lastly, each sound condition was presented without visual stimuli (2 conditions) resulting in 11 conditions in all.

The two sound conditions were dropped due to potential interference from the noise of the scanner. This left 3 conditions, face, body and face and body together (figure). All 20 clips and their variations i.e. 60 clips in total were digitised and the task was computerised to allow use in an fMRI paradigm and furthermore to allow the measurement of reaction times.

Two additional conditions were designed to control for brain activation due to subsidiary processes:

1) button press condition - a 'stop/go' task simply requiring a button press in response to 'go'. This is necessary for cognitive subtraction of activation due to simply viewing the screen and pressing the button.

2) 'non-mental state' inference condition - participants have to decide on the physical 'channel' displaying the mental state from one of two of the following options i.e. body, face or figure. As this condition uses the same video clips as those used in

the experimental task, this condition controls for activation as a result of viewing the video clips and making a decision based upon them using a button press.

The final task had 180 clips 60 blank, 60 non-mental state inference – 20 face, 20 body, 20 figure and 60 mental state attribution – 20 face, 20 body, 20 figure. Participants were required to choose one of two response alternatives displayed on the screen. Each clip lasts approximately 2 secs although the response choices were displayed for 8 secs (see Figure 7.1).

Figure 7.1: Cycle for each stimuli in the PONS-r

NO IMAGE WORDS	VIDEO CLIP WORDS	NO IMAGE WORDS	NO IMAGE NO WORDS

0-----------1.5--------------------3.5------------------------7.5-------------8*

1.5 secs	2 secs	4 secs	.5 secs

*words disappear once choice has been made and at 8 secs cycle for next stimuli commences

2.2. Mental state labels

In the original study the correct and incorrect answers from which the selection was made were often presented as short sentences. For the revised version the answers and their foils were updated and made more succinct by reducing each mental state descriptor to a maximum of 2 words. Four colleagues were sent the 'figure' clip, the original word and alternative suggestions and were asked to either endorse one of the given suggestions or provide an alternative. The final labels were chosen according to the majority endorsement. This meant that no label was chosen without at least 2 people in agreement although in the majority of cases all 4 raters and the experimenter agreed. The original and final labels can be found in appendix 7.

2.3 Counterbalancing

2.3.1 Foils: experimental

In the original version of the task, allocation of response alternatives were random with adjustments made so that no clip was over- or under-represented. However, as the original task had a different number of conditions, the order of presentation was reworked for the current paradigm and the similarity of the response labels were equally balanced rather than randomly allocated as this variable is integral to task difficulty. Across non-mental state inferences i.e. for face, body and figure – 5 foils were similar mental states[25], 5 were the same broad valence (i.e. positive or negative) and 5 were the opposite valence.

2.3.2 Foils: control

Again this was balanced so that each possibility appeared an equal amount of times. For example, when 'face' was the correct alternative – the incorrect alternative was:

5 body left

5 body right

5 figure left

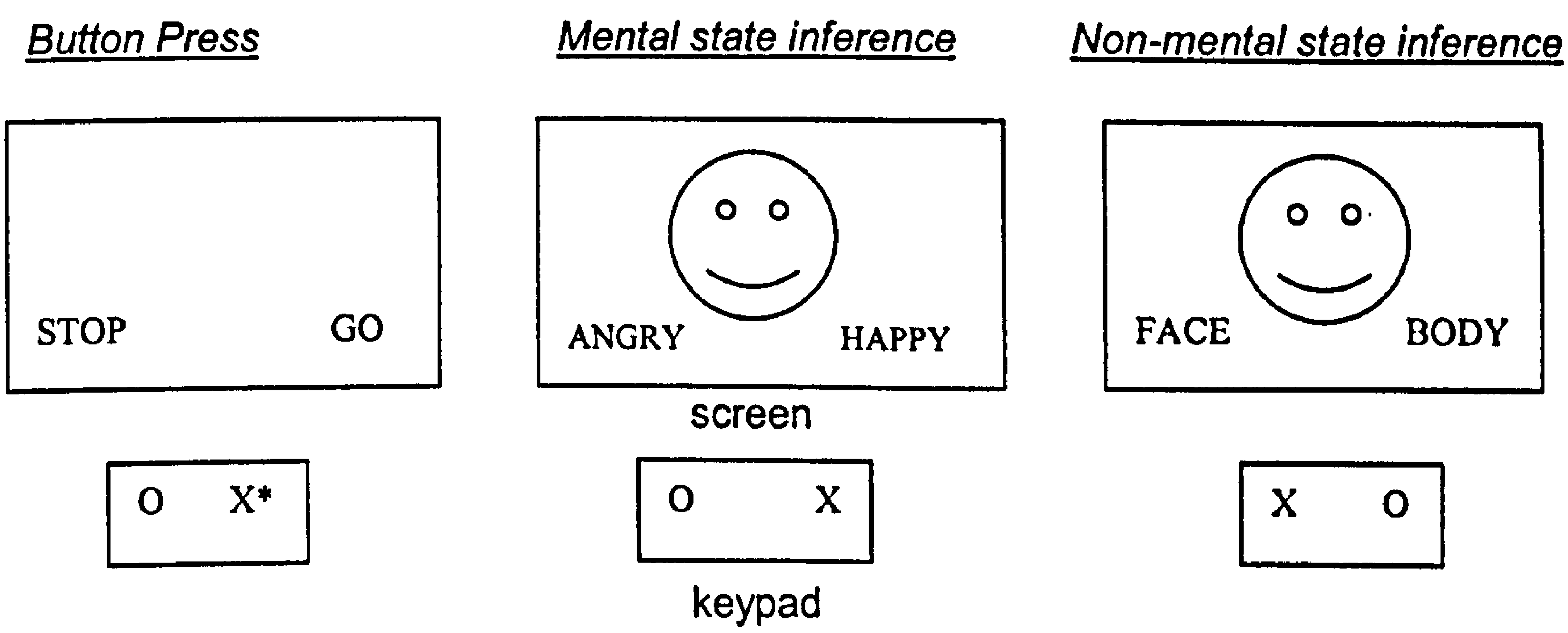
5 figure right

[25] Based on the four original categories of positive dominant, positive submissive, negative dominant and negative submissive.

2.3.3. Button press condition

'Go' appears half the time on the right and half on the left.

Figure 7.2: Example of stimuli in each condition#



each stimuli runs through the cycle detailed in Figure 1.
* X = correct answer and O = incorrect answer.

2.3.4. Randomisation Checks

All the variables were correlated with the clip number to see if there was a tendency for any one condition, response alternative or physical 'channel' to be overrepresented at the beginning or end of the task. The only significant variable was 'face' in the mental state inference condition meaning that there was a slight tendency for face only clips in this condition to appear at the beginning – however this was a low correlation ($r = -.273$, $n = 20$, $p = .035$) and this didn't hold for both mental and non-mental state inference conditions combined.

3. Offline Pilot Study

3.1 Design

The online PONS-r has 3 conditions (see Figure 7.2 above):

1) button press condition - a 'stop/go' task simply requiring a button press in response to 'go'.

2) non-mental state inference condition - participants have to decide on the physical 'channel' displayed from one of two options i.e. body, face or figure.

3) experimental condition - participants are shown clips of 20 mental states, one at a time and have to choose the correct answer after each, from 2 alternatives.

The clips are graded in difficulty in two ways; 1) the non-mental state inference displayed – some clips just show the actors body alone, which does not have as many cues as the face or figure, hence making it more difficult to decipher 2) the similarity of the response alternative – task difficulty is increased if the incorrect answer is very similar to the correct one.

3.2 Procedure

The instructions were altered to fit the new paradigm and were as follows:

Please choose the word that best describes the behaviour or feelings of the person in the video clip.

Initially, participants were given an Information Sheet and the opportunity to ask any questions. They were then given a practice run consisting of 9 items and were allowed to do this as many times as they wanted. The error rate and reaction times were measured for all conditions. Next they were given the task proper along with the 'Reading the Mind from the Eyes task' (Baron-Cohen *et al.*, 2001 - see chapter 6), the EQ (Baron-Cohen and Wheelwright, 2004 in press) and the IRI (Davis, 1980 - see chapter 2).

3.3 Participants

Nine men and 11 women recruited from the Institute of Psychiatry (11) and local area (9). Their mean age was 35.1 years (± 10.43).

3.4 Results

The mean and standard deviation scores can be found in table 7.1.

Table 7.1 Mean accuracy rate on the PONS-r out of 60

	Mental State Inference	Non-Mental State Inference	Button Press
Mean	43.7	57.3	59.5
SD	2.74	1.69	.69
n	20	20	20

The expected pattern was observed with the mental state items invoking more errors, followed by the 'non-mental state inference' items and the lastly the button press condition. The ~~error~~ accuracy rate on the mental state items was remarkably similar to the normative data (72.8 % in the current study as opposed to 77% in the original datasets) and was high enough to be above chance and low enough to indicate a lack of ceiling effects. The range of scores was, however, small which is probably due to the limited sample size.

In terms of the two different types of 'difficulty' manipulations i) lexical – similarity of the foil and ii) non-mental state inference of presentation, mean accuracy rates were computed (see Table 7.2). The expected pattern of results was observed with the highest accuracy rate on stimuli were the face or figure was shown, dropping when participants were just shown the 'body' items. The pattern for lexical difficulty however deviated slightly from the predicted pattern as it appeared that the medium items were slightly easier than the 'easy' items, however, the accuracy rate may have been elevated by the increased spread of scores as a result of having double the amount of items.

Table 7.2 Mean accuracy rate on the each condition of the PONS-r

	Lexical 'easy' items - out of 15	Lexical 'medium' items - out of 30	Lexical 'difficult' items - out of 15	Figure - mental state items - out of 20	Body - mental state items - out of 20	Face - mental state items - out of 20
Mean	10.5	24.7	9.4	15.8	11.6	16.05
SD	1.15	1.95	1.47	1.4	1.47	1.61
N	20	20	20	20	20	20

An item by item inspection revealed that across physical channels all items except 3 were endorsed as correct by more than 50% of sample, and those with a higher error rate only occurred when the response alternative was very similar i.e. thankful vs. affection, and/or when the 'physical' channel was body.

In terms of validity, some modest, but not statistically significant, correlations were observed between performance on the 'Reading the Mind in the Eyes task' and accuracy on the body items (r = .407, n = 20, p = .075) which may reflect the fact that both tasks are similar inasmuch as they are designed to be difficult being based on limited cues. A positive association was also observed with between performance on body items and 'personal distress' scores on the IRI ($r = .392$, $p = .088$) consistent with the findings from previous chapters (see chapter 4) that an egocentric affective response to other's emotional states is related to accurate mental state labelling.

4. Online study: Methods

4.1 Participants

An fMRI study was conducted with 12 right-handed healthy volunteers. There were 6 men and 6 women and the mean age was 32.2 years (± 9.95). None of the participants had a history of mental or neurological illness nor did they have any contraindications for scanning. The mean verbal IQ as predicted from the NART was 119.4 (± 5.9). All participants were also pre-screened with the EQ and scored within normal range on the EQ for their gender (see chapter 2).

4.2 Procedure

Prior to the study, participants were given standardised instructions, a consent form and screened for MRI suitability. They also completed the Eyes Test (Baron-Cohen *et al.*, 2001).

4.3. Image acquisition

Participants were then scanned using GE Signa 1.5T Neuro-optimised MR system (General Electric, Milwaukee, WI, USA). After the collection of high resolution structural scans (3.3mm thick) in the axial plane, an event-related design was employed with 60 clips in each of the 3 conditions presented in a randomised order, therefore, totalling 180 clips. With a TR of 2 secs and each of 180 stimuli lasting 8 seconds (video clips 2 secs and ISI 6 secs), 720 gradient echo echoplanar T2* weighted images were acquired at each of 16 near-axial, 7mm thick planes. Accuracy rate and reaction times were also measured; these data were not used in the statistical analysis but simply to check that the participant followed the instructions and that the constraints of the MRI scan did not adversely affect performance on the task.

4.5 Image analysis

The data were analysed using in-house software (XBAMv2 and 3), a non-parametric method of analysis which derives probabilities of significance from random permutations. Images were first corrected for head motion (Bullmore *et al.*, 1999). Next, the physiological response and corresponding change in the T_2 weighted signal intensity as a result of the experimental conditions, was estimated by using a least squares fitting of a sinusoidal regression model at each voxel for all the images (Bullmore, 1996). The regression model included a pair of sine and cosine waves to allow the shape of the haemodynamic response to be estimated at each voxel.

Next the fundamental power quotient (FPQ – fundamental power divided by its standard error) was estimated at each voxel and represented in a descriptive parametric map which was registered in standard stereotactic space (Talairach and Tournoux, 1988) and smoothed by a 2D Gaussian filter (± 5mm). Permutation testing was then used to identify voxels demonstrating significant median standardised power of response (median FPQ), over all subjects, during each task with one-tailed probability of false positive error set at $p = .001$ This process was randomly permuted 10 times and the FPQ estimated after each permutation (Brammer *et al.*, 1997).

In summary, the permutation procedure consisted of each time series in the observed image being randomly permuted and the sinusoidal regression model fitted by generalised least-squares after permutation. This gave an estimate of standardised power under the null hypothesis that the estimated FPQ is not determined by the experimental design. Repeating this procedure 10 times at each voxel, generated 10 maps under the null hypothesis for each participant. Each map was transformed into standard space and smoothed, just as the observed maps are. Lastly, the null distribution of the median FPQ was computed, by taking each participant's median FPQ at each voxel for each set of permuted FPQ maps. The median FPQ observed at each voxel was then tested at the critical value corresponding to the one-tailed probability of a type 1 error, $p = .001$. Voxels that 'pass' this critical threshold were considered 'activated' specific to the experimental design. Such a non-parametric approach is thought to be particularly appropriate for small sample sizes where outliers can adversely affect the dataset.

In addition, each participants dataset was manually checked for movement and as 3 participants were shown to have some movement artefact, these datasets were registered using a longer, more thorough registration process in xbam_v3. These datasets were also detrended as it is possible that other variables can influence data acquisition which can be removed by detrending using different filters.

5. Online study: results

5.1 Behavioural

T-tests were conducted to examine any group differences on task performance between those who performed the task online and those that undertook the task outside of the scanner. Group differences were observed on the 'figure' items only ($t = 6.074$, df 30, $p < .001$). Examination of the means show that the 'online' group performed significantly less well with a mean score of 12.6 as opposed to 15.8 in the offline group.

It was reasoned that this difference was probably due to the stimuli images being viewed at a further distance in the online task due to restrictions imposed by fMRI scanning. It was not possible to blow up the images to compensate due to their poor resolution. It would be expected that this difference between on- and offline presentations would exert its biggest influence in the 'figure' which is a smaller image (see appendix 8). The magnitude of the difference between the online and offline condition indicates a slight, although significant reduction in overall accuracy.

The relationship between other empathy measures (Eyes test and EQ) and accuracy rates on the PONS-r was then examined for the entire group (those who did the task both offline and online). Items of medium lexical difficulty were positively related to 'emotional reactivity' as measured by the EQ ($r = .339$, df 31, $p = .062$). The relationship between 'body' items and 'social skills' as measured by the EQ ($r = .314$, df 31, $p = .084$) was also approaching significance.

5.2 fMRI Data

In order to explore any differences between condition in mean power of functional activation, generic brain activation maps (GBAMs) were constructed. The subtraction technique was employed to isolate the area of activation specific to the experimental conditions.

5.2.1 Simple Analyses

In order to examine brain activation specific to the experimental condition, first activation specific to 'non-mental state inference' > control was considered. The non-mental state inference task used a forced choice format to prompt participants to decide whether the channel shown was body, face or figure. Activation in this condition relates primarily to viewing the images and making a non-mental state inference as opposed to simply reading some words and pressing a button in response (i.e. the stop/go control task).

Areas of brain activation are reported in Table 7.3 along with voxel size and the maximum FPQ (see above). The main areas of activation were seen in the precentral gyrus, visual areas including V1 and V2 and parastriate areas, anterior mid cingulate, frontal pole, angular gyrus, fusiform gyrus, middle temporal gyrus, medial frontal and dorsolateral prefrontal cortex, amongst other areas

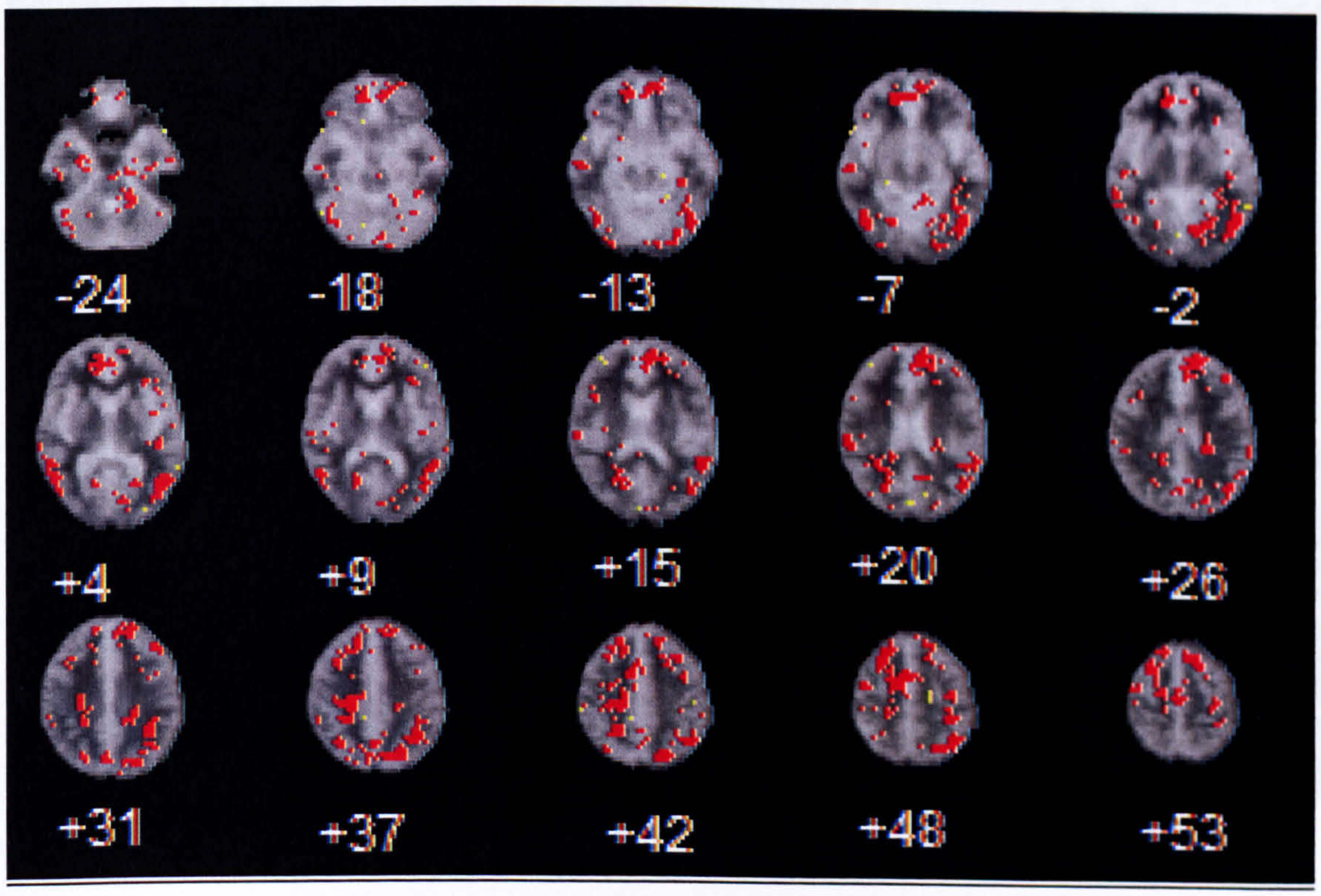
Table 7.3: Brain areas activated for non-mental state inference > button press[26]

No.	x	y	z	Side	BA	Brain Region	Max FPQ
50	15	-23	53	R	4	Precentral Gyrus (M1)	0.050943
32	-25	-69	31	L	19	Primary Visual Cortex	0.049853
29	11	-17	37	R	24	Ant-Mid Cingulate Gyrus	0.052118
25	-17	46	-7	L	10	Frontal Pole	0.050745
24	-21	-60	37	L	7	Precuneus	0.058946
24	-11	-76	26	L	19	Primary Visual Cortex	0.047918
23	-47	-46	-2	L	21	Middle Temporal Gyrus	0.063561
21	-36	-46	-7	L	37	Fusiform Gyrus	0.041057
20	-17	-80	-7	L	18	Primary Visual Cortex	0.047169
19	25	13	53	R	6	Premotor Cortex & SMA	0.066316
19	-11	43	15	L	32	Medial Frontal Lobe	0.099519
19	-36	-37	37	L	40	Supramarginal Gyrus	0.044399
18	15	30	37	R	9	DLPFC	0.038271
16	-32	-26	48	L	4	Precentral Gyrus (M1)	0.047641
9	-43	-60	26	L	39	Angular Gyrus	0.043621

Brain activation was extensive as the 'button press' condition does not control for visual input (see Figure 7.3). However, this comparison allows the identification of areas specific to the 'non-mental state inference'. It is also possible that participants may have 'implicitly' decoded the mental states in this condition (see discussion for further explanation).

[26] This analysis is hypothesis driven and only areas relating to the experimental hypotheses are reported. All areas are activated at p = .001.

Figure 7.3: 'Non-mental state inference > button press group brain activation maps[1]



The next comparison was between areas activated for the 'mental state inference' task over and above that observed in the 'non-mental state inference' task. There are less areas of activation here due to the matching of control conditions (see Figure 7.4).

[1] Left is right - red is 'non-mental state inference' and yellow is 'button press' - activation at a probability threshold of 0.001 is shown. Numbers are distance in mm above or below the AC-PC line.

Figure 7.4: Mental state inference > non-mental state inference group brain activation maps[1]

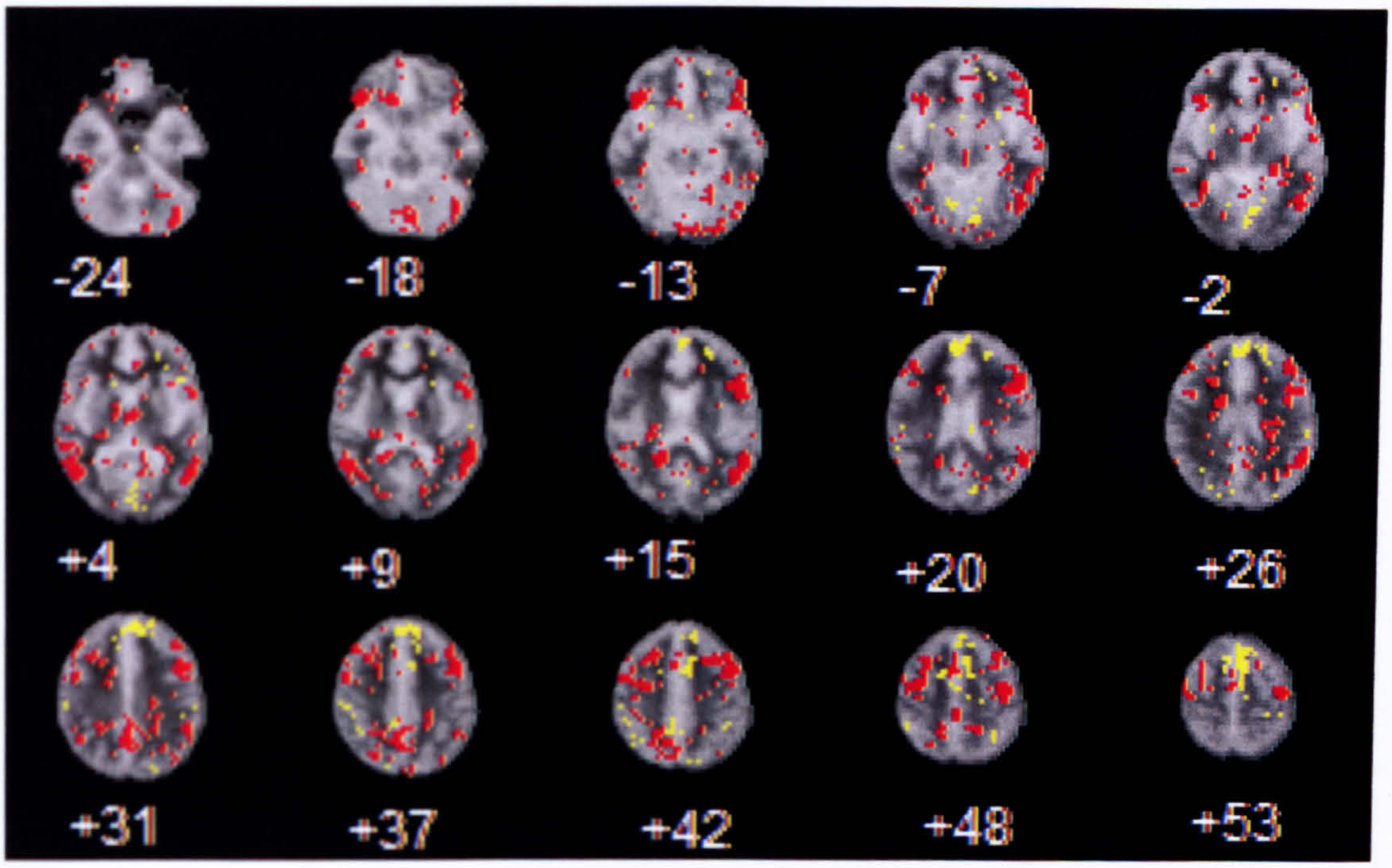


Table 7.4 gives further details of this comparison. Activation specific to the 'mental state inference' task was observed in areas such as the dorso- and ventro-lateral prefrontal cortex, medial frontal lobe and superior temporal gyrus, inferior frontal gyrus, premotor cortex, supramarginal gyrus, precentral gyrus, insula.

[1] Again, left = right but here red is the 'mental state inference' and yellow 'non-mental state inference.

Table 7.4: Areas of brain activation for mental state inference >non-mental state inference

No.	x	y	z	Side	BA	Brain Region	Max FPQ
29	- 40	13	42	L	9	DLPFC	0.064775
29	4	-37	37	R	31	Post Cingulate Gyrus	0.072134
25	- 43	23	20	L	45	Inf Frontal Gyrus (Broca's)	0.082364
23	-53	-33	9	L	22	Superior Temporal Gyrus	0.086157
23	-61	-37	-7	L	21	Middle Temporal Gyrus	0.069781
18	-47	30	-13	L	47	VLPFC	0.063568
17	25	20	37	R	32	Medial Frontal Lobe	0.068854 ~~16~~
16	-53	-39	26	L	40	Supramarginal Gyrus	0.069683
16	-40	-10	48	L	4	Precentral Gyrus (M1)	0.092911
14	4	13	48	R	6	Premotor Cortex & SMA	0.083023
13	-53	-50	-2	L	37	Middle Temporal Gyrus	0.082817
10	-43	20	-18	L	38	Ant Temporal Pole	0.079585
10	36	20	-2	R	72	Insula	0.079585
7	15	-39	48	R	5	Superior Parietal Lobule	0.069841

5.2.2 Channel analyses

In order to explore areas of activation specific to decoding mental states from the body alone, the data were analysed to allow comparisons between the physical 'channel' of presentation (see Table 7.5). By examining these contrasts together, it is possible to circumnavigate the potential confound created by the 'body' stimuli being more difficult to decipher than the 'face' and 'figure' stimuli. Although, there are a lack of data regarding the neural substrates of emotion perception from the body, it is expected that the medial areas of the fusiform gyrus be activated in line with Hadjikhani and DeGelder (2003).

Table 7.5: Activation specific to 'body in the mental state inference condition'*

No.	x	y	z	BA	Brain Region
Face > Body					
10	0	-74	-1	18	Primary Visual (Peristriate V2, V3)
Body > Face					
9	43	-64	5	37	R Middle Occipital gyrus
8	0	-57	37	18	Primary Visual (Peristriate V2, V3)
Body > Figure					
14	-45	-66	-2	19	L Primary Visual Cortex
6	50	24	22	47	R Inferior Frontal Gyrus
Figure > Body					
8	-15	0	32	24	L Anterior Mid Cingulate Gyrus
6	-14	-13	42	23	L Posterior Cingulate Gyrus
Figure > Face					
18	-16	-1	34	20	L Inferior Temporal Gyrus
18	29	4	50	6	R Premotor Cortex & SMA
13	-27	13	45	6	L Premotor Cortex & SMA
Face > Figure					
8	46	28	22	16	R DLPFC

*Data reported here have a high number of expected false positives in comparison to those observed and hence carry a risk of type 1 errors, but as activations are theoretically plausible and relevant to our hypotheses, they are reported nevertheless.
NB Figure = face and body

Activation specific to face > body was observed in visual areas. Activation for body > face was observed in the middle occipital gyrus and visual areas (see Figure 7.5a). Activation specific to body > figure was again observed in visual areas and the inferior frontal gyrus (see Fig 7.5b). Activation relating to figure > body, however was observed in the anterior cingulate (see Fig 7.5c). Whereas activation relating to figure > face was observed in the inferior temporal gyrus (Fig 7.5d). By a process of deduction, it is reasonable to suggest that the areas relating to body are indeed areas invoked when reasoning with particularly difficult material whereas the figure vs. body comparison (when considered in light of the fig vs. face activation) relates to activation specific to body without the added confound of difficulty (see below).

Fig 7.5a: Body > Face[1]

Middle occipital gyrus

Fig 7.5b: Body vs. Fig

Inferior frontal gyrus

[1] Left is right for Fig 7.5 a, b, c, d. Slices illustrated show the greatest contrast between conditions.

Fig 7.5c: Fig > Body

Anterior Mid Cingulate Gyrus

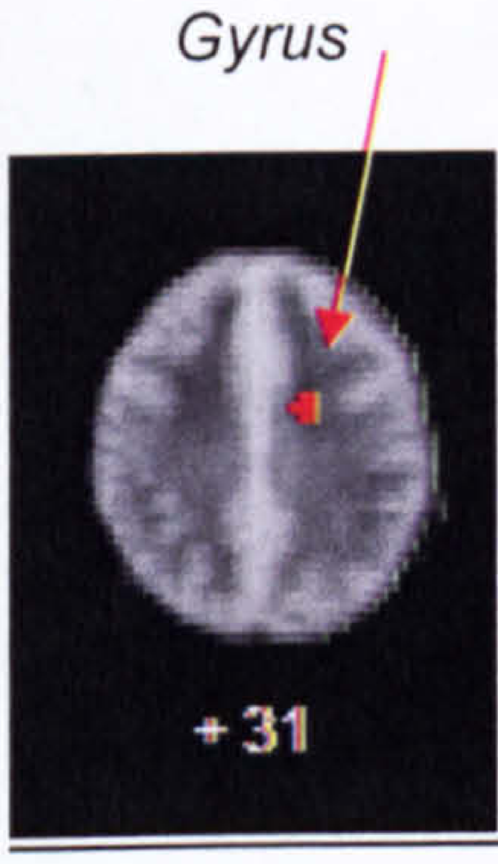
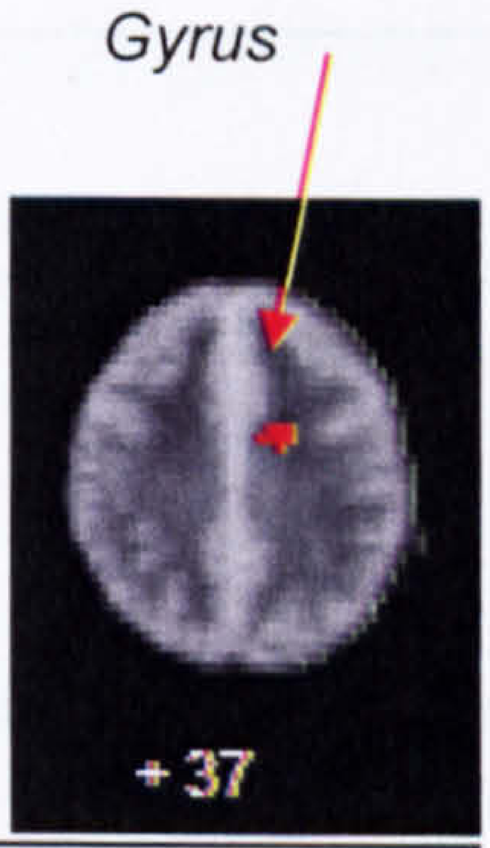


Fig 7.5d: Fig > Face

Inferior Temporal Gyrus



5.2.3 Lexical analyses

No differences in activation were observed between conditions which were lexically 'easy' i.e. the foil was dissimilar to the correct answer, or lexically difficult i.e. the foil was very similar to the correct answer, suggesting the same strategies were at play irrespective of this aspect of task 'difficulty'.

5.2.4. Behavioural measures

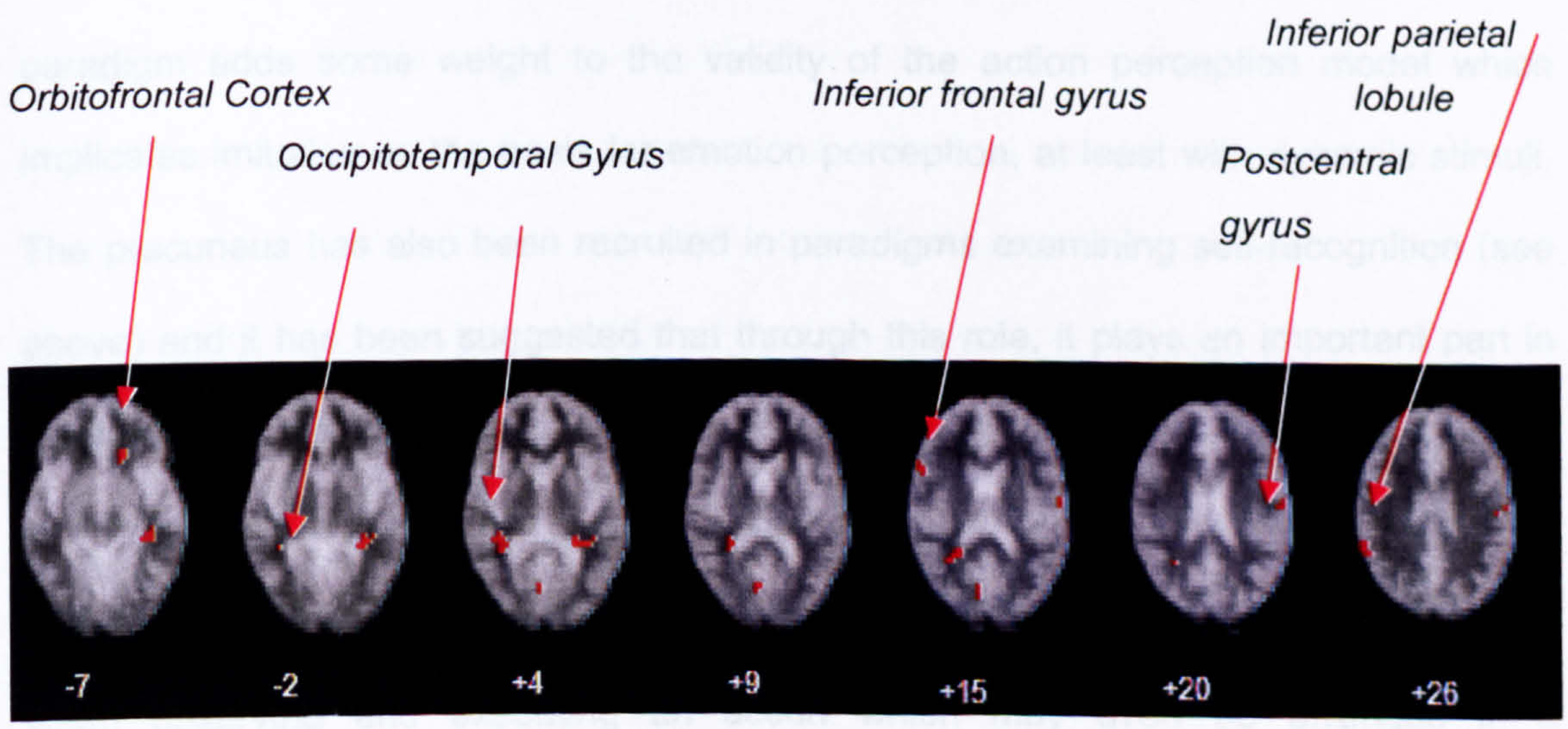
As there was an association between 'emotional reactivity' and items of medium lexical difficulty i.e. 50% of the mental state attribution items, these behavioural data were correlated with activation from the mental state inference task > non-mental state inference task (see Table 7.6).

Table 7.6: Brain areas that showed a positive association with 'emotional reactivity' scores

No.	x	y	z	BA	Brain Region
20	35	-35	-3	36	R Occipitotemporal Gyrus
15	-16	27	-12	11	L Orbitofrontal Cortex
14	32	-39	5	36	R Occipitotemporal Gyrus
6	-57	-10	19	21	L postcentral gyrus
7	55	- 41	27	37	R Inferior Parietal Lobule
4	52	14	14	47	R Inferior frontal gyrus

Areas that were positively correlated with 'emotional reactivity' levels include the orbitofrontal cortex, occipitotemporal gyrus (fusiform face area), inferior frontal gyrus, postcentral gyrus and the inferior parietal lobule (see Figure 7.6).

Figure 7.6: Group brain activation maps for areas that correlate with 'emotional reactivity'



This suggests these areas may be central to mental state attribution in those with an increased tendency to experience an affective response to others' emotional states.

6. Discussion

The main findings from this study were that brain areas typically associated with the use of 'shared representations' (Decety and Chaminade, 2003) to understand other people's mental states were activated alongside those thought to underlie traditional ToM reasoning. This is consistent with a theory of mental state attribution based on a model of action recognition (Decety, 2002) as expected from a task tapping the neural correlates of dynamic stimuli. In addition, these findings are in line with the ideas of Adolphs *et al.*, (2000) that 'covert online somatosensory representations' may be of particular use when the stimuli to be decoded are especially difficult.

The pre-motor cortex and SMA, precentral gyrus and precuneus were also activated in a series of studies by Decety and colleagues investigating action recognition and imitation (see above). To find these areas activated in an emotion perception paradigm adds some weight to the validity of the action perception model which implicates imitation as the basis for emotion perception, at least with dynamic stimuli. The precuneus has also been recruited in paradigms examining self-recognition (see above) and it has been suggested that through this role, it plays an important part in discriminating self from other.

In addition, the pre-motor cortex is thought to contain mirror neurones that fire both when observing and executing an action which may even be arranged in a somatotopic manner (Buccino *et al.*, 2001). The precentral gyrus (M1) was also activated during a study examining the first and third person simulation of action and so is thought to represent another area of overlap between self and other representations (Ruby and Decety, 2001). The fact that it was activated here again suggests that this area is involved in representing others actions, including actions depicting the others' mental states.

A series of studies have suggested the superior parietal lobe is vital to distinguishing self from other (see above). To find it activated here, in the mental state inference condition, is therefore consistent with previous findings. The same is true of the observed activation of the posterior cingulate. Emerging evidence suggests this structure to be involved primarily in higher order representation of the self (and by this route it may be involved in representations about other people) as it has been activated in paradigms tapping self representation.

Somatosensory areas such as the insula and supramarginal gyrus were also activated in the current study. Activation of the supramarginal gyrus is consistent with the idea that the inferior parietal lobule plays an important role in distinguishing self from other. Adolphs *et al* (2000) also found both these areas to be important for emotion perception, especially when the target mental state was difficult to decode from static stimuli. It seems that mental states that are particularly difficult to decode may require the construction of a 'covert online somatosensory' representation. In addition, the insula is thought to be involved in disgust (Phillips *et al.*, 1998), both its experience and perception (Wicker *et al.*, 2003a). Carr *et al.*, (2003) also found the insula to be implicated both when participants observe and imitate the emotional expressions of others. The activation of these areas in the current paradigm is no surprise as the current stimuli was both dynamic and difficult.

The activation of Broca's area is of particular interest due to the recent suggestion that this area be involved in imitation and perhaps contains mirror neurones being adjacent to premotor areas (but see Ramnani and Miall, 2004 for discussion). This raises interesting questions regarding the link between language and mental state attribution and for the understanding of some impairments in clinical groups such as those with Asperger's Syndrome (see earlier chapters). However, interpretation of this finding must be pursued with caution as it is possible that the extra verbal load in this mental state inference task could account for it. An additional explanation is that people attempted to lip read in some of the conditions or had a tendency for 'silent speech' although participants debriefing suggests this was not the case.

As expected, areas associated with traditional ToM task were also recruited in this paradigm, consistent with previous fMRI work. These include the DLPFC and medial frontal cortex and medial temporal gyrus. The medial frontal cortex and medial temporal gyrus are involved in ToM attribution, according to numerous studies (see above). The DLPFC has also been found to be involved in empathy, and it is thought it may exert its influence through cognitive flexibility (Eslinger, 1998). Alternatively, others have also found the DLPFC to be involved in the preparation of action and it is entirely possible that its role in empathic ability is based upon this. Furthermore, the medial superior temporal gyrus which is thought to be specialised in perceiving facial affect, in particular from the eye region, was also recruited.

More complex comparisons were conducted to examine activation specific to the perception of mental states from the body vs. face area and difficult vs. easy stimuli. No differences in activation were found at a probability level low enough to eliminate false positives, indicative of a lack of differential activation. The lack of difference in activation between figure and body in the areas relating to 'shared representations' goes some way to suggest that with dynamic stimuli at least, an action perception model is appropriate regardless of task difficulty. However, there is also the possibility that the 'figure' items were actually more difficult in this paradigm by virtue of them being hard to see (see above). But the lack of differential inactivation for body > face also suggests that 'difficult' stimuli i.e. body, did not invoke special circuits. In addition, there was no difference in activation between items that were considered harder in terms of lexical difficulty i.e. easy vs. hard. It seems therefore that, at least with dynamic stimuli, a simulation approach may be appropriate independent of task difficulty, in contrast to the differential processing observed between difficult and easy static stimuli (Adolphs *et al.*, 2000).

However, despite this, some differential activation did exist at a probability of 0.05. Although, these results carry a high risk of a type 1 error, as they fit well with the literature they require some discussion. The inferior frontal gyrus, ~~close to Broca's area~~ was activated in the body > figure. This comparison is most likely to tap task 'difficulty' and this finding is consistent with the use of 'mirror neurones' in decoding other people's mental states. However, perhaps this result is not surprising when one considers that the 'body' stimuli were in one sense composed solely of gestures. This may provide one explanation for the evolutionary basis of 'mirror neurones' in an area traditionally specialised for speech. The middle occipital lobe was also activated in the body > face condition, and this area has been shown to be activated when perceiving 'biological motion' from point light displays (Vaina *et al.*, 2001) and the fusiform gyrus was also activated in a study looking at static bodily expressions of fear (Hadjikhani and DeGelder, 2003). This therefore suggests that this area is being activated in response to decoding mental states from bodily expressions rather than in response to difficult stimuli.

The existence of these subtle differences in activation combined with the lack of global difference suggests that some key areas may be activated more when additional simulative processing is needed. In other words, the same strategy is employed with dynamic stimuli regardless of the task difficulty, but activation is stronger in certain key areas when the task is particularly difficult in an 'additive' manner. These differences could therefore to some extent be viewed as quantitative rather than qualitative. This is consistent with Adolph's analogy (see above) regarding the different resources required to visualise the different aspects of your own house. If this is the case, then these data suggest that ~~Broca's area and~~ the use of mirror neurones, are key to the recognition of mental states especially when the task requires full use of the available resources.

The general lack of differences does not support the functional specialisation of brain areas for particular body parts i.e. face vs. body, at least with dynamic stimuli, as elicited in this paradigm. However, again, the differential activation observed when the probability threshold was lowered may shed some light on the processing of mental state information from different body areas. The 'figure' clips are most ideally placed for examining the neural correlates of decoding mental states from bodily expressions, as they are not confounded by 'difficulty' issues to the same degree as the 'body' clips. Figure > body activated the anterior mid cingulate and posterior cingulate both of which have been shown to be involved in representing information regarding the self (Kelley *et al.*, 2002) including bodily information (Lane *et al.*, 1997, Critchley *et al.*, 2003, Critchley *et al.*, 2001). Of course, it could be argued that this comparison merely represents the inclusion of a face in the 'figure' stimuli but activation observed in the face > figure comparison does not support this explanation. The inferior temporal gyrus was also activated in figure > face, as were premotor areas, perhaps reflecting weak somatotopic activation. It seems therefore possible to conclude tentatively that the same areas are involved in decoding mental states from different parts of the body, but that the anterior, posterior cingulate and inferior temporal gyrus may play a particular role in decoding bodily expressions (ie., without the face), in addition to the middle occipital lobe (see above).

Behavioural measures were also correlated to examine their association with different areas of brain activation. Scores on the 'emotional reactivity' component of the EQ were found to be significantly correlated with activation in the 'fusiform face area', orbitofrontal cortex, inferior frontal gyrus, postcentral gyrus and inferior parietal lobule. The inferior parietal lobule is thought to be a central component of the 'shared representations' network playing a key role in distinguishing the 'self' from 'other'. The inferior frontal gyrus is thought to contain mirror neurones (see above) and perhaps this activation suggests those high on 'emotional reactivity' are more likely to utilise the representations produced by such neurones. To find increased activation in the fusiform gyrus, an area thought to be specialised for perceiving faces, in people high on 'emotional reactivity', is also of interest. In addition, the orbitofrontal cortex is thought to be involved in the more complex aspects of empathy such as integrating epistemic and affective states (Eslinger, 1998, Damasio, 1994). Activation in the postcentral gyrus fits well with a simulation account of 'empathy' (see above) and so it is of interest that this area is more activated in those who self-report overt affective responses in response to other people's mental states.

These data suggest that such areas are fundamental to attributing mental states in those who rate themselves as high on 'emotional reactivity' and are therefore more likely to respond affectively to others' mental states. Future work should attempt to distinguish between 'affective empathy' and 'personal distress' since 'emotional reactivity' may tap both concepts (see chapter 2). However, it is also possible that these two types of emotional reactivity share some cognitive processes and hence similar brain activation patterns (see chapter 4). Adams (2001) points out that a person feeling despair in response to another person's despair, may have their 'mirror neurones firing' even if the despair is not felt for the target and the response is one of 'personal distress'.

No amygdala activation was observed in this study. However, a methodological criticism of fMRI and other techniques reliant on the subtraction method may have a bearing on this finding. Areas that are activated in both comparison conditions may be 'cancelled out'. Activation presented as specific to an experimental condition is actually activation specific to that condition over and above a comparison condition. This can lead to important areas of overlapping activation being missed. This issue may be particularly problematic in the current study as 'implicit' mental state attribution may have occurred in the 'non-mental state inference condition' (see Carr *et al.*, 2003). Unfortunately, this was unavoidable due to the need to control for the activation of the visual areas. Although the choice alternatives were displayed for 1.5 secs before the clip in an attempt to circumnavigate this effect and prime people to attend to the correct aspect of the clip, the overlap between activation in both mental state and non-mental state inference conditions suggests implicit mental state processing probably still occurred. Although this does not detract from the validity of the findings for the mental state inference task, some important areas of activation may have been missed. Having said this, the similarity in activation between the two conditions suggests that areas were *more* activated in the condition requiring explicit mental state attribution as they remained significantly activated when contrasted with the non-mental state inference task (see Carr *et al.*, 2003). However, this does mean that it is not possible to say conclusively that the amygdala was not activated.

Further areas that may have been 'subtracted' out include the precentral gyrus; 50 voxels were activated above threshold in the 'non-mental state inference' condition and only 16 in the 'mental state inference' condition. The lack of activation in the 'fusiform gyrus' in the main analysis may well also be a result of such 'cancelling' out. The fact that it was associated with 'emotional reactivity' suggests this may be the case. However, this finding is also consistent with the idea that this area is actually specialised for 'expertise' as opposed to faces *per se* (Haxby *et al.*, 2002). The angular gyrus was also activated in the non-mental state condition but not in the 'mental state condition' and it is possible that this area is important for mental state appreciation. This area has been found to be activated during reflection on one's own personality traits (Kjaer *et al.*, 2002) and electrical stimulation of this area has induced the sensation of an out-of-body experience (Blanke *et al.*, 2002). In addition, lesions in the dominant angular gyrus have been associated with body image disturbances. These data therefore suggest this area plays a key role in representing the self, which is also consistent with previous observations of activation in the inferior parietal lobule being vital for distinguishing self from other.

There are few limitations in this study that require discussion and clarification. A potential confound is having seen constituent 'parts' of the video clips before as body and face are again shown in the figure condition. This means that for up to 2/3 (although probably less depending on the order of presentation) of the simulation conditions it is technically possible to work out the right answer from memory by remembering which response alternative appeared before either in a figure or face/body clip. However to adopt such a strategy would be quite effortful as one would have to remember seeing the face/body/figure and corresponding words, all in a very short space of time. Previous studies using the PONS have not controlled for this potential confound and participant debriefing indicated that this strategy was not adopted.

Further limitations general to fMRI imaging include the reliability of methods used to localise function. Swallow *et al.,* (2003) argue that group mapping may be unreliable, in part be due to the fact that anatomical differences are treated as random error. It seems that although fruitful to examine groups of data in order to learn where a function is localised, as in the current study, to examine the individual generic brain activation maps may be more reliable. It may therefore be worth re-examining the current datasets individually at a later date.

Lastly, the fact that this study used a whole-brain approach carries certain limitations. The brain areas reported were selected on the basis a priori hypotheses and voxel size, however, much activation went unreported which may mean key areas are missed. This is especially true when attempting to localise relatively new process such as this aspect of empathy, where there are only a handful of previous studies of direct relevance to generate hypotheses. With such cautions in mind, the data from this study does seem to support the utility of an action recognition model of emotion perception particularly when using dynamic stimuli under non-optimal conditions.

6.1 Conclusion

It seems that an action perception model based on the use of 'shared representations' is of use for decoding other people's mental states when presented using dynamic stimuli and the main nodes in the neural network supporting are the inferior frontal gyrus, premotor cortex and SMA, precentral gyrus, supramarginal gyrus, insula, DLPFC, precuneus and other areas in the inferior and superior parietal lobules. However, more information is needed on the conditions in which this strategy is adopted. The data presented here tentatively suggests that this approach is of use with all dynamic stimuli rather than that which are simply difficult to decode.

The relationship between the processes recruited for both dynamic and static mental state processing remains unclear. Are the areas typically thought to be involved in emotion perception from static stimuli still key to perceiving those same states from dynamic stimuli? The data presented here complements neuropsychological (Humphreys *et al.*, 1993) and neuroimaging studies (LaBar *et al.*, 2003) which suggest not. However, further investigation is necessary to tease apart any possible dissociation. To conclude, although it seems likely that the attribution of mental states from dynamic stimuli is reliant on 'shared representations', more information as to whether there are particular conditions that encourage this strategy.

6.2 Further Work

Pilot studies using the offline PONS-r are already underway with the clinical groups discussed elsewhere in this thesis as a prelude to online studies. The prediction is that people with DPD may have particular problems with this kind of emotion perception, as to be expected from a group whose sense of self (a vital component of using 'shared representations') is in disarray. People with AS, however, may perform better on some aspects of this task than they do in other areas of social cognition, consistent with their performance on the tasks described in chapter 5.

Discussion

The cognitive neuropsychiatric approach is based on the application of the principles of cognitive psychology and neuropsychology to the understanding of psychiatric conditions and behavioural phenomena (Halligan and David, 2001). This involves developing a cognitive model of normal function and then applying it in the clinical and neuropsychological domain. The aim of this thesis was to explore empathy, and in so doing construct a normal model of empathic processes that will be of value in explaining clinical phenomena. Two conditions that result in, or from, a disruption of empathy were given special attention: Asperger's Syndrome and Depersonalisation Disorder.

Contents at a glance

1. Empathy conceptualised

As a concept empathy is relatively new in English-speaking countries, although the term 'sympathy' was in use much earlier and often to denote what we now call empathy (Darwin, 1872). Social psychologists have been the most prolific of the empirical researchers on this subject, providing a distinction between cognitive and affective empathy early on (Hogan, 1969, Mehrabian and Epstein, 1972). If we attempt to define these concepts, a further debate becomes relevant: whether there are separate processes specialised for perceiving others' epistemic vs. affective states, or 'hot' vs. 'cold' cognition (Stone *et al.*, 2003), which also translates loosely (though not completely) into 'socio-perceptual' and 'socio-cognitive' theory of mind' – ToM - (Tager-Flusberg and Sullivan, 2000).

A further concept important in the study of empathy has been proposed: 'personal distress' (Davis, 1980). Although this is not empathy *per se*, it does resemble empathy inasmuch as it is an emotional response to others' affective states. This similarity has often led to conceptual confusion. A model incorporating this distinction together with new ideas from the clinical domain, i.e. ToM was detailed in Chapter 1.

Several self-report measures appear to uphold the distinction between cognitive and affective empathy, and those that have measured 'personal distress' also suggest its importance. However, all these measures have been shown to have important weaknesses, and until recently none were designed with the clinical condition in mind. As a result, the first empirical chapter of this thesis (chapter 2) sought to clarify and confirm the cognitive and affective distinction, using a new self-report measure, the Empathy Quotient (EQ - Baron-Cohen and Wheelwright, 2004 ~~in press~~), and statistical methods. In so doing it further validated the EQ. Confirmation of this came when an exploratory factor analysis uncovered 3 main factors: i) 'cognitive empathy', ii) 'emotional reactivity', and iii) 'social skills', consistent with traditional ideas of empathy. Unfortunately, items did not distinguish between 'personal distress' and 'affective empathy' – hence the assignment of the broader label for factor 2.

2. Empathic processes

As much previous work has focused on the perception and detection of epistemic states, under the program of research aimed at detailing the ToM, we explored the perception of affective states from complex stimuli. In the past, emotion attribution has only been studied mainly in the context of simple paradigms, for the most part based on labelling facial emotional expression (Hadjikhani and DeGelder, 2003). The detection of affective states is more characteristic of empathy, as is shown by everyday usage of the term, plus the fact that epistemic state detection is usually related to cognitive empathy alone. On the other hand, affective states can be the focus of both cognitive and affective empathy. In addition, we reasoned in chapter 4 that the two key components of the empathic process need to be captured empirically: the cognitive representation and physiological response.

In designing a task to tap the cognitive representations that form the basis of empathy, ideas were drawn from the literature on philosophy of mind and psychology, where the process of simulation has been proposed as being of relevance to the perception of affective states (Gallese and Goldman, 1998). Recent findings, ideas and models in neuroscience also support this claim. We designed and developed a paradigm using one little-known study by a social psychologist exploring the use of the self-representation in perspective-taking as a guide (see chapter 4 and Davis *et al.*, 1996). Data from a control group of 53 participants showed that, as predicted, self-other overlap was associated with a person's self-reported 'emotional reactivity' in response to other's mental state. These findings are consistent with a 'simulation' account of the appreciation of affective states. Further evidence of the role of the 'self' in the perception of others' affective states was gathered from the study of clinical groups (see below).

In addition, an implicit task designed to tap physiological arousal was used. Although the concept of speech-rate has been used within psychology before (Teasdale and Fogarty, 1979, and see chapter 4), this is the first time it has been used in emotion detection research. It has an important advantage over overt methods of measuring arousal such as those based on subjective report inasmuch as it is implicit, and hence not prone to demand characteristics. The control data gathered from the use of this task suggests it may assist in measuring physiological arousal.

To develop a cognitive test of both of these key components of empathy was essential to build a useful model of empathy. This was especially necessary as the recent resurgence in the study of empathy and related processes has been based largely on neuroscientific rather than on cognitive data (see chapter 7). However, these tests only scratch the surface since many unanswered questions remain. First, what is simulation: is it a process specialised for affect perception, or a domain general process relying on skills such as IQ? Secondly, is simulation the process of choice for the detection of all affective states (and indeed some think epistemic states), or is it only called upon during particular circumstances? Thirdly, does physiological arousal occur when we perceive all affective states, and is the difference between cognitive and affective empathy as applied to affective states one of degree that is to say, covert vs. overt physiological arousal? These questions should form the basis of future work.

3. Cognitive function and empathy

The clinical groups detailed in this thesis were shown to have various problems with the different components of empathy (see below). However, this was not related to verbal intelligence or general cognitive functioning, as both groups were shown to do particular well on different cognitive tests (and even surpass the performance of the control group).

One such task involved conditional reasoning which is a form of logical deductive reasoning (Kemp *et al.*, 1997). For different but predictable reasons both clinical groups outperformed control participants on this task (see chapter 3). This suggests that these 'clinical' conditions might be better viewed as cognitive styles rather than disabilities (Baron-Cohen, 2002b). It also indicates that this type of reasoning is distinct from empathic processes (this, however, is only a single dissociation and should be viewed cautiously – see below).

Participants were also tested on a physics task (Baron-Cohen *et al.*, 1999b), and again performed in line with controls. Interestingly, an inverse association was observed between reasoning in the physical and empathic domains, suggesting there may be some kind of trade-off between these two processes. These data again speak against empathy being based on general reasoning processes, and, more importantly for this thesis, confirm that the clinical data detailed below were not simply due to a failure of general processes.

4. Empathy and Asperger's Syndrome

People with ASDs are often viewed by clinicians and laypeople as lacking in empathy. To a certain degree this is true. Much research suggests that people with ASDs lack the ability to comprehend representational or epistemic states (see chapter 1 and 6). People with ASDs are often impaired in appreciating others' mental states, including affective states, that rarely occur in isolation. These skills are usually integral to sociability. However, affective empathy in people with ASDs has been to some extent overshadowed by research into traditional ToM. Despite this, anecdotal reports and indirect findings indicated that any impairments in social cognition may be restricted to representational ToM (Blair, 1999, Frith, 2003). The offline tasks developed in this thesis were ideally placed for testing this possibility. With regard to physiological arousal as measured by speech rate, people with Asperger's Syndrome displayed the same patterns as controls: congruent physiological arousal in response to the affective states of the protagonist. Despite hints that this might be the case, these data is still surprising, as the notion that people with ASDs are 'emotionally cold' is far-reaching. This finding was also consistent with some of the self-report measures in which people with Asperger's Syndrome did not score differently to controls i.e. on 'empathic concern' items. Their lower scores on 'emotional reactivity' seemed to be due to the inclusion of items tapping meta-representation rather than an innate lack of empathy.

In addition, the scores on the trait-overlap task were also to some extent surprising. They suggest that the people with ASDs used an online strategy to appreciate other people's affective states: utilising their own self-concept to understand others as opposed to relying on compensatory strategies and 'wooden algorithms' (Zahavi and Parnas, 2003). But the data also suggests that the strategy was anchored in egocentric rather than allocentric processes (Langdon and Coltheart, 2001), suggesting that this skill may be underdeveloped in people with ASDs. However, this is a large improvement on the ability of people with AS to detect epistemic states, which is at best described as severely disrupted, any apparent success being based on alternative strategies (see below and chapter 6).

To find spared, or at the very least partially preserved, skills in the social domain in people with ASDs, is potentially exciting, all the more so against the backdrop of deficits and inadequacies observed in this area. However, on two counts these results need to be interpreted cautiously: scientifically, the data need to be replicated and the methods refined; and in terms of ecological validity, because affective states are rarely presented in isolation from representational states, and the fractionation of social cues for empirical investigation does not reflect the constant flow of information in the real world. For all this, to find skills that are potentially spared in this domain, may help in interventions designed to enable people with ASDs to navigate through the social world. Future work should examine whether these findings are specific to Asperger's Syndrome as opposed to all autistic spectrum disorders, especially as the brain areas that subserve language and simulation may overlap.

5. Empathy and Depersonalisation Disorder

People with Depersonalisation Disorder (DPD) often report a subjective lack of empathy. Until now, however, these reports remained anecdotal. A few studies examined emotion processing in people with DPD and found evidence of the 'emotional numbing' often reported by this group (Phillips *et al.*, 2001, Sierra *et al.*, 2002). But no studies specifically examined the perception of other people's emotions in DPD sufferers'. Paradoxically, the data from the self-report scales (see Chapters 2 and 3) suggest that this group do not have problems with affective empathy. However, evidence of emotional blunting was found in the participants' responses to the 'personal distress' items of the Interpersonal Reactivity Index (Davis, 1980). And furthermore, the objective empathy measures were consistent to some extent with the anecdotal reports.

The results of the speech rate measure of physiological arousal indicate that as a group, people with DPD do not experience congruent physiological arousal consistent with an empathic response. In fact the speech rate pattern displayed by this group showed the opposite profile to the AS and control groups. This is no surprise when one considers the anecdotal reports of lack of empathy, along with the empirical and subjective reports of 'emotional blunting'. It is intuitive that a lack of emotional arousal concerning the 'self' would transfer to emotional arousal regarding 'others'. This therefore suggests that an intact sense of personal emotional experience is necessary to enjoy emotional resonance with other people.

Furthermore, the tendency of some people with DPD to perform less well on labelling 'basic' emotional states suggests that in certain circumstances physiological arousal may also be needed to simply identify or perhaps bolster the identification of others' emotional states. However, this result was not significant, and only occurred in a handful of participants. But this is the kind of question we need to answer in order to further our understanding of emotion perception, empathy and the processes that are key. The fact that the AS and control groups showed congruent physiological arousal in this task also hints at a vital role for emotional experience in emotional perception *per se* (including cognitive empathy) as opposed to affective empathy. Although the tasks were designed to be ecologically valid, they were still part of an experimental manipulation, and so to find physiological arousal under these conditions suggests the qualitative aspects of an emotion may have a general role in emotion detection.

The DPD group were also shown to use the self-concept in apprehending other people's emotional states, though again the overlap was in excess of that observed in the control group (and more so than in the AS group). As with the AS group, these data was interpreted in line with an egocentric/allocentric distinction. The most plausible explanation is that an excess of trait overlap is indicative of egocentric simulation (Langdon and Coltheart, 2001). To find the DPD group to be over-reliant on the self-concept in this domain is in line with recent cognitive-behavioural (CBT) models of DPD which propose an excessive focus on the self. This possible causal factor in a lack of subjective empathy is open to immediate empirical exploration as newly emerging CBT models attempt to defocus the DPD client from the self. It would be of interest, therefore, to see if this strategy had any effect on reinstating the emotional experience and in turn the empathic one. This may be of value to DPD clients who no doubt experience additional distress due to this disruption in the empathic process.

The fact that both those with AS and DPD showed the same excess of trait overlap suggests the model lacks specificity. More work is need to explore this issue and to plot the differences and similarities between these two groups' who appear phenomenologically distinct. The self-other overlap needs to be fully operationalised and factors such as the optimum levels explored. Using the 'double dissociation' method (Shallice, 1989) with these groups, who at first glance appear similar on this variable, may help further fractionate this skill. Data presented here from these two distinct groups have already been used to this end to demonstrate how some components of empathy are dissociable.

6. Clinical associations and dissociations.

In this thesis, people with DPD and those with AS displayed dissimilar speech rate patterns indicative of differential physiological arousal in response to the affective states of others. That people with DPD were found to perform well on tasks designed to tap ToM cognitive empathy (see chapter 6), despite their lack of physiological arousal, suggests that these two processes are functionally dissociable. This explanation is further strengthened by the fact that people with AS only tended to do well on the ToM tasks by using compensatory processes, suggesting they were in fact impaired. This hints at a 'double dissociation' of function between these two components of empathy.

To find the affective and cognitive components of empathy dissociated in this way is of use in localising and fully understanding the processes at work in empathy. Furthermore, such 'dissociation' may extend further than between physiological arousal and ToM reasoning alone. That people with AS had a congruent physiological response after perceiving others' affective states suggests they understood those states in a qualitative manner. That they fail to appreciate qualitatively epistemic states, and that the reverse pattern is found in those with DPD, suggests the apprehension of epistemic and affective states or 'hot' and 'cold' cognition relies on different cognitive (and perhaps neural) processes. This finding has important consequences not only for different clinical conditions, but also for the eventual localisation of these processes, and the development of a cognitive model.

7. The neural basis of empathy

The 'offline' data we have gathered suggest that a simulation approach may be of use in apprehending affective states. Furthermore, the data from those with DPD go some way to reinstate the role of emotional experience in emotion perception. In addition, the dissociations of function proposed suggest that perhaps the system subserving traditional ToM, with its onus on epistemic states, differs from that specialised for affective states. This is consistent with the emerging literature that suggests that 'action perception' (Blakemore and Decety, 2001b) and the use of 'shared representations' (Decety and Chaminade, 2003) may provide a useful framework for 'hot' cognition (see Chapters 1 and 7).

This idea has been around for a while, and empirically the groundwork has been laid in the form of studies examining the neural correlates of action perception vs. generation, and various manipulations such as imitation. However, it is only very recently that actual emotion recognition or mental state attribution paradigms have been studied experimentally within this framework (for instance see Decety and Chaminade, 2003 and chapter 7, Carr *et al.*, 2003).

In this thesis, stimuli from an existing social psychology experiment were used (Rosenthal *et al.*, 1979), in order to explore the application of action recognition models to mental state attribution. The stimuli had two helpful components that made them appropriate for exploring the link between these two models: it was dynamic and difficult (Adolphs, 2001a). Both of these factors are thought to induce 'shared representations' as the strategy of choice. The stimuli were revised to suit an fMRI paradigm, and validated offline (see chapter 7).

Activation specific to the experimental condition was consistent with the use of 'shared representations' in decoding other people's mental states. It included brain areas that are thought to house 'mirror neurones' (Broca's Area and premotor cortex), and areas key to distinguishing the self from other (parietal lobe) and those central to constructing online representations (somatosensory areas such as the insula). This data therefore suggests that an 'action perception' model is suited to explaining mental state attribution, especially when the stimuli are dynamic (which is often the case) and can be viewed as emotion action.

Further data are needed to establish whether such an approach is limited to dynamic stimuli, or under what conditions it may be of use with static stimuli. To find brain activation consistent with the use of 'shared representations, which is very much allied to 'simulation', provides further support for the utility of this approach. To find converging evidence from two different levels of explanation is encouraging, and an important part of the 'falsification' process (Popper, 1963).

8. Limitations

This thesis had various general limitations. The clinical groups were incompletely characterised on some of the measures, for example, the full-scale IRI (Davis, 1980), because the decision to administer the full scale was taken *post hoc.* But the data from these measures were interpreted with caution and in conjunction with other findings. There was also some missing data that added to this incomplete characterisation i.e. the speech rate. However, as power calculations were based on stringent levels (95%) enough statistical power remained to detect difference on the basis of the existing data. In addition, due to time constraints some of the tasks were not administered in their entirety, for example, the Rowe 'theory of mind' tasks, which may have limited the conclusions draw from these data. Nevertheless, the decision was made in light of the available time and factors such as participants concentration and tiredness.

Many of the tasks used we used were exploratory and will need refinement and replication. This is especially so for the tasks detailed in chapter 4. In addition, the data from the clinical groups were not replicated, and hence conclusions need to be tentative. It could also be argued that the thesis included a lack of standardised measures. However, attempts were made to complement the use of exploratory tasks, by administering standardised traditional ToM tasks (see chapter 6).

9. Future Research

The data gathered in this thesis generates new hypotheses and predictions. First, and specifically, the model detailed in Chapter 1 needs to be fully operationalised and empirically tested. Secondly, the novel and objective tests of empathy need to be refined and further validated. Thirdly, the nature of simulation and similar processes needs to be explored, and a cognitive model generated with greater specificity. These steps would in turn enable a more thorough exploration of the empathic process in the clinical groups detailed. The next logical step which is already underway, is to carry out functional neuroimaging on clinical groups detailed in this thesis whilst doing the PONS-r, the results of which will add to the cognitive data reported in chapter 5 and hence strengthen our knowledge regarding these conditions.

10. Conclusions

Summarised, the main findings from this thesis are: i) it is possible to quantify empathy using a combination of subjective and implicit measures; ii) that people with AS have some elements of empathy preserved; iii) people with DPD have trouble responding with affective empathy; iv) 'simulation' and 'shared representations' are central to empathy and have a distinct functional neuroanatomy. Many unanswered questions remain regarding the nature of these processes and their disruption. However, it is hoped that ~~the fact that~~ the findings in this thesis are of relevance to at least two clinical conditions, and also that they help specify a normal neuropsychological model of empathy. If so the utility of a cognitive neuropsychiatric approach will have been bolstered.

Bibliography

Abdel-Khalek, A., Lester, D. and Barrett, P. (2002) The factorial structure of the Arabic Obsessive-Compulsive Scale in Kuwaiti and American college students *Personality & Individual Differences,* **33,** 3-9.

Adams, F. (2001) Empathy, neural imaging and the theory versus simulation debate *Mind & Language,* **16,** 368-392.

Adolphs, R. (1999) Social cognition and the human brain *Trends in Cognitive Sciences,* **3,** 469-479.

Adolphs, R. (2001a) Neural systems subserving emotion: lesion studies of the amygdala, somatosensory cortices, and ventromedial prefrontal cortices. In *Handbook of Neuropsychology*, Vol. 5 (Eds) Gainotti, G. Elselvier Science.

Adolphs, R. (2001b) The neurobiology of social cognition *Current Opinion in Neurobiology.,* **11,** 231-9.

Adolphs, R. (2002) Recognising emotion from facial expressions: Psychological and neurological mechanisms *Behavioral and Cognitive Neuroscience,* **1,** 21-62.

Adolphs, R. (2003) Cognitive neuroscience of human social behaviour *Nature Reviews Neuroscience,* **4,** 165-178.

Adolphs, R., Baron-Cohen, S. and Tranel, D. (2002) Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience* **14,** 1264-1274.

Adolphs, R., Damasio, H., Tranel, D., Cooper, G. and Damasio, A. R. (2000) A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping *Journal of Neuroscience* **20,** 2683-2690.

Adolphs, R., Damasio, H., Tranel, D. and Damasio, A. R. (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* **372,** 669-672.

Adolphs, R., Russell, J. A. and Tranel, D. (1999) A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli *Psychological Science* **10,** 167-171.

Adolphs, R., Sears, L. and Piven, J. (2001a) Abnormal processing of social information from faces in autism *Journal of Cognitive Neuroscience.* **13,** 232-40.

Adolphs, R. and Tranel, D. (2003) Amygdala damage impairs emotion recognition from scenes only when they contain facial expressions *Neuropsychologia* **41,** 1281 - 1289.

Adolphs, R., Tranel, D. and Damasio, A. R. (2003) Dissociable neural systems for recognising emotions *Brain and Cognition* **52,** 61-69.

Adolphs, R., Tranel, D. and Damasio, H. (2001b) Emotion recognition from faces and prosody following temporal lobectomy *Neuropsychology* **15(3),** 396-404.

Amaral, D. and Corbett, B. (2003) The amygdala, autism and anxiety. In *Autism: Neural basis and treatment possibilities.* (Eds) Bock, G. and Goode, J. John Wiley and Sons Ltd, Chichester.

Avikainen, S., Kulomaki, T. and Hari, R. (1999) Normal movement reading in Asperger subjects *Neuroreport,* **10,** 3467-70.

Avikainen, S., Wohlschlager, A., Liuhanen, S., Hanninen, R. and Hari, R. (2003) Impaired mirror-image imitation in asperger and high functioning autistic subjects *Current Biology,* **13.**

Bach, L., Happe, F., Fleminger, S. and Powell, J. (2000) Theory of mind: independance of executive function and the role of the frontal cortex in acquired brain injury. *Cognitive Neuropsychiatry,* **5,** 175-192.

Bacon, A., Fein, D., Morris, R. G., Waterhouse, L. and Allen, D. (1998) The responses of autistic children to the distress of others. *Journal of Autism & Developmental Disorders,* **28,** 129-142.

Baker, D., Hunter, E., Lawrence, E., Medford, N., Patel, M., Senior, C., Sierra, M., Lambert, M. V., Phillips, M. L. and David, A. S. (2003) Depersonalisation disorder: clinical features of 204 cases *British Journal of Psychiatry.,* **182,** 428-33.

Baron-Cohen, S. (1989) Are autistic children behaviourists? An examination of their mental-physical and appearance-reality distinctions *Journal of Autism & Developmental Disorders,* **19,** 579-600.

Baron-Cohen, S. (1991) Do people with autism understand what causes emotion? *Child Development* **62,** 385-395.

Baron-Cohen, S. (1993) The development of theory of mind, and its dysfunction.In *Understanding Other Minds: Perspectives from Autism* (Eds).Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D.

Baron-Cohen, S. (1999a) The extreme male-brain theory of autism.In *Neurodevelopmental disorders. Developmental cognitive neuroscience* (Ed)Tager-Flusberg, H., pp. 401-429.

Baron-Cohen, S. (1999b) Theory of mind and autism: a fifteen year review.In *Understanding Other Minds: Perspectives from Developmental Neuroscience* (Eds).Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D.

Baron-Cohen, S. (2002a) The Extreme Male Brain theory of autism *Trends in Cognitive Sciences* **6,** 248-254.

Baron-Cohen, S. (2002b) Is Asperger Syndrome necessarily a disability. *Focus on Autism* **17.**

Baron-Cohen, S. and Hammer, J. (1997) Is autism an extreme form of the "male brain"? *Advances in Infancy Research* **11,** 193-217.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R. and Plaisted, K. (1999a) Recognition of faux pas by normally developing children and children with Asperger's Syndrome or high functioning autism *Journal of Autism & Developmental Disorders* **29,** 407 - 418.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N. and Wheelwright, S. (2003) The Systemising Quotient: An investigation of adults with Asperger's syndrome or high functioning autism, and normal sex differences *Philosophical Transactions of the Royal Society; Biological Sciences* **358,** 361-374.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C. and Williams, S. C. R. (2000) The amygdala theory of autism *Neuroscience & Biobehavioral Reviews* **24,** 355-364.

Baron-Cohen, S. and Wheelwright, S. (2004 ~~in press~~) The Empathy Quotient: An investigation of adults with Asperger's Syndrome or high functioning autism, and normal sex differences *Journal of Autism & Developmental Disorders,* 34(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. and Plumb, I. (2001) The "Reading the Mind in the Eyes" Test Revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism *Journal of Child Psychology & Psychiatry & Allied Disciplines* **42,** 241-251.

Baron-Cohen, S., Wheelwright, S., Stone, V. and Rutherford, M. (1999b) A mathematician, a physicist and a computer scientist with Asperger syndrome: Performance on folk psychology and folk physics tests *Neurocase* **5(6) 1999.**

Baron-Cohen, S., Wheelwright, S., Stott, C., Bolton, P. and Goodyer, I. (1997) Is there a link between engineering and autism? *Autism* **1**, 153-163.

Barrett, J. and Paus, T. (2002) Affect-induced changes in speech production *Experimental Brain Research* **146**, 531-537.

Batson, C. D., Fultz, J. and Schoenrade, P. A. (1987) Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences *Journal of Personality* **55**, 19-39.

Beck, A. T., Epstein, N. and Brown, G. (1988a) An inventory for measuring clinical anxiety: Psychometric properties *Journal of Consulting and Clinical Psychology* **56**, 893-897.

Beck, A. T., Steer, R. A. and Garbin, M. G. (1988b) Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation *Clinical Psychology Review* **8**, 77-100.

Belmonte, M. and Yurgelun-Todd, D. (2003) Functional anatomy of impaired selective attention and compensatory processing in autism *Cognitive Brain Research* **17**, 651-664.

Bernstein, E. M. and Putnam, F. W. (1986) Development, reliability, and validity of a dissociation scale *Journal of Nervous & Mental Disease* **174**, 727-735.

Beversdorf, D., Anderson, J., Manning, S., Anderson, J., Nordgren, R., Felopulos, G., Heilman, K. and Bauman, M. (1998) The effect of semantic and emotional context on written recall for verbal language in high functioning adults with autism spectrum disorder. *Journal of Neurology, Neurosurgery & Psychiatry* **65,** 685-692.

Blackburn, S. (1996) *The Oxford Dictionary of Philosophy,* Oxford University Press, Oxford.

Blackshaw, A., Kinderman, P., Hare, D. and Hatton, C. (2001) Theory of mind, causal attribution and paranoia in asperger syndrome *Autism* **5,** 147-163.

Blair, J. (1999) Psychophysiologial responsiveness to the distress of others in children with autism *Personality & Individual Differences* **26,** 477- 485.

Blair, J. (2003a) Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society; Biological Sciences* **358,** 561-572.

Blair, J. (2003b) Neurobiological basis of psychopathy *British Journal of Psychiatry* **182,** 5-7.

Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., & Jones, L. (1996) Theory of mind in the psychopath *Journal of Forensic Psychiatry* **7,** 15-25.

Blair, R. J. R. (2003c) Facial expressions, communication and neural substrates *Philosophical Transactions of the Royal Society; Biological Sciences* **358,** 561-572.

Blake, R., Turner, L., Smoski, M., Pozdol, S. and Stone, W. (2003) Visual recognition of biological motion is impaired in children with autism *Psychological Science* **14,** 151-157.

Blakemore, S. J. (2003a) The detection of contingency and animacy from simple animations in the human brain., *Cerebral Cortex* **13,** 837-844.

Blakemore, S. J. (2003b) Self-awareness and action, *Current Opinion in Neurobiology* **13,** 219-224.

Blakemore, S. J. and Decety, J. (2001a) From the perception of action to the understanding of intention, *Nature Neuroscience Reviews* **2,** 561-7.

Blakemore, S. J. and Decety, J. (2001b) From the perception of action to the understanding of intention, *Nature Reviews Neuroscience* **2,** 561-7.

Blakemore, S. J., Oakley, D. and Frith, C. D. (2003) Delusions of alien control in the normal brain, *Neuropsychologia* **41,** 1058 - 1067.

Blakemore, S. J., Wolpert, D. M. and Frith, C. D. (1998b) Central cancellation of self-produced tickle sensation., *Nature Neuroscience* **1,** 635-640.

Blanke, O., Ortigue, S., Landis, T. and Seeck, M. (2002) Stimulating illusory own-body perceptions, *Nature.* **419,** 269-70.

Bormann-Kischkel, C., Vilsmeier, M. and Baude, B. (1995) The development of emotional concepts in autism, *Journal of Child Psychology & Psychiatry & Allied Disciplines* **36,** 1243-1259.

Brammer, M., Bullmore, E., Simmons, A., Williams, S., Grasby, P., Howards, R., Woodruff, P. W. and Rabe-Hesketh, S. (1997) Generic brain activation mapping in functional magnetic resonance imaging: a non-parametric approach. *Magnetic Resonance Imaging* **15,** 763-770.

Breitenstein, C., Van Lancker, D. and Daum, I. (2001) The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample *Cognition and Emotion* **15,** 57-79.

Brentano (1895) *Psychology from an empirical standpoint,* Routledge, London.

Brierley, B., Shaw, P. and David, A. S. (2004 ~~in press~~) Emotional memory and perception in temporal lobectomy patients with amygdala damage *Journal of Neurology, Neurosurgery & Psychiatry*, 75 (4), 593 - 599.

Brothers, L. and Ring, B. (1992) A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience* **4,** 107-118.

Buccino, G., Binkofski, F., Fink, G., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R., Zilles, K., Rizzolatti, G. and Freund, H. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience* **13,** 400-404.

Buitelaar, J. K. W., M (1997) Are deficits in the encoding of affective cues and in mentalising abilities independent *Journal of Autism and Developmental Disorders* **27,** 539-556.

Bullmore, E. (1996) Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance Medicine* **35,** 261 - 277.

Bullmore, E., Brammer, M., Rabe-Hesketh, S., Curtis, V. A., Morris, R. G., Williams, S. C. R., Sharma, T., Murray, R. and McGuire, P. K. (1999) Methods for the diagnosis and treatment of stimulus correlated motion in generic brain activation studies using fMRI. *Human Brain Mapping* **7,** 38 - 48.

Cacioppo, J., Martzke, J., Petty, R. and Tassinary, L. (1988) Specific forms of facial EMG response index emotions during interview: From Darwin to the continuous flow hypothesis of affect-laden information processing. *Journal of Personality and Social Psychology* **54,** 592-604.

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H. and Gabrieli, J. D. (2002) Amygdala response to happy faces as a function of extraversion *Science.* **296,** 2191.

Capps, L., Yirmiya, N. and Sigman, M. (1992) Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology & Psychiatry & Allied Disciplines* **33,** 1169-1182.

Carlson, E. B. and Putnam, F. W. (1993) An update on the dissociative experiences scale *Dissociation: Progress in the Dissociative Disorders* **6,** 16-27.

Carlson, E. B., Putnam, F. W., Anderson, G., Clark, P., Torem, M., Coons, P., Bowman, E., Chu, J., Dill, D., Loewenstein, R. J. and Braun, B. G. (1991) *Factor analysis of the dissociative experiences scale; a multi-center study,* Rush Presbyterian, Chicago.

Carr, L., Iacoboni, M., Dubeau, M., Mazzlotta, J. and Lenzi, G. (2003) Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas *Proceedings of the National Academy of Sciences of the United States of America.* **100,** 5497-5502.

Castelli, F., Frith, C., Happe, F. and Frith, U. (2002) Autism, Asperger Syndrome and brain mechanisms for the attribution of mental states to animated shapes *Brain* **125,** 1839-1849.

Cattell, R. B. (1966) The scree test for the number of factors *Multivariate Behavioural Research* **1,** 140-161.

Chaminade, T., Meltzoff, A. N. and Decety, J. (2002) Does the end justify the means? A PET exploration of the mechanisms involved in human imitation *Neuroimage.* **15,** 318-28.

Channon, S., Charman, T., Heap, J., Crawford, S. and Rios, P. (2001) Real-life-type problem-solving in Asperger's Syndrome *Journal of Autism & Developmental Disorders.* **31,** 461-9.

Charman, T. (2002) Understanding the imitation deficit in autism may lead to a more specific model of autism as an empathy disorder *Brain and Behavioural Sciences* **25,** 29-30.

Charman, T. and Baron-Cohen, S. (1994) Another look at imitation in autism *Development & Psychopathology* **6,** 403-413.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G. and Drew, A. (1997) Infants with autism: an investigation of empathy, pretend play, joint attention and imitation *Developmental Psychology* **33,** 781-789.

Cook, N. D. (2003) Tone of voice and mind: The connections between intonation, emotion, cognition and consciousness *Advances in Consciousness Research* **47,** 97-99.

Corona, R., Dissanayake, C., Arbelle, S., Wellington, P. and Sigman, M. (1998) Is affect aversive to young children with autism? Behavioural and cardiac responses to experimenter distress *Child Development* **69,** 1494-1502.

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C. R., Van Amelsvoort, T., Robertson, D. M., Rowe, A., Phillips, M., McAlonan, G., Howlin, P. and Murphy, D. G. M. (2000) The functional neuroanatomy of social behaviour: Changes in cerebral blood flow when people with autistic disorder process facial expressions *Brain* **123(11),** 2203-2212.

Critchley, H. D., Mathias, C., Josephs, O., O'Doherty, Zanini, S., Dewar, B., Cipolotti, L., Shallice, T. and Dolan, R. (2003) Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence *Brain* **126,** 2139-2152.

Critchley, H. D., Melmed, R., Featherstone, E., Mathias, C. and Dolan, R. (2001) Brain activity during biofeedback relaxation *Brain* **124,** 1003- 1012.

Crowne, D. P. and Marlowe, D. (1960) A new scale of social desirability independent of psychopathology *Journal of Consulting Psychology,* **24,** 349-354.

Dailey, M. N., Cottrell, G. W., Padgett, C. and Adolphs, R. (2002) EMPATH: A neural network that categorizes facial expressions *Journal of Cognitive Neuroscience* **14,** 1158-1173.

Dalgleish, T. and Power, M. J. (Eds.) (1999) *Handbook of cognition and emotion,* (1999). xxi, 843pp.

Damasio, A. (1994) *Descartes' error: Emotion, reason, and the human brain,* Avon Books, New York.

Darwin, C. (1872) *The expression of the emotions in man and animals,* (1872).

Davies, M. and Stone, T. (1995) *Mental simulation - Evaluations and applications,* Blackwell, Oxford.

Davies, S., Bishop, D., Manstead, A. and Tantam, D. (1994) Face perception in children with autism and Asperger's Syndrome *Journal of Child Psychology & Psychiatry & Allied Disciplines* **35,** 1033-1057.

Davis, M. (1980) A multidimensional approach to individual differences in empathy *Catalog of Selected Documents in Psychology* **10,** 85.

Davis, M., Conklin, L., Smith, A. and Luce, C. (1996) Effect of perspective-taking on the cognitive representation of persons: A merging of self and other *Journal of Personality and Social Psychology* **70,** 713-726.

Davis, M., Hull, J., Young, R. and Warren, G. (1987) Emotional reactions to dramatic film stimuli: The influence of cognitive and emotional empathy *Journal of Personality and Social Psychology* **52,** 126-133.

Davis, M. H. (1994) *Empathy: A social psychological approach,* Brown & Benchmark, Dubuque IA.

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T. and Meyer, M. (1999) Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors *Journal of Personality* **67**, 469-503.

Decety, J. (2002) Neurophysiological evidence for simulation of action.In *Simulation and Knowledge of Action* (Eds).Dokic, J. and Proust, J. Benjamins Publishers, Philadelphia, pp. 53-72.

Decety, J. and Chaminade, T. (2003) Neural correlates of feeling sympathy *Neuropsychologia* **41**, 127-138.

Decety, J., Chaminade, T., Grezes, J. and Meltzoff, A. N. (2002) A PET exploration of the neural mechanisms involved in reciprocal imitation *Neuroimage.* **15**, 265-72.

Decety, J. and Sommerville, J. (2003) Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences* **7**, 527-533.

Dennis, M., Lockyer, L. and Lazenby, A. (2000) How high-functioning children with autism understand real and deceptive emotion *Autism* **4**, 370-381.

Dimberg, U., Thunberg, M. and Grunedal, S. (2002) Facial reactions to emotional stimuli: Automatically controlled emotional responses *Cognition & Emotion* **16**, 449-472.

Doherty, R. (1998) Emotional Contagion and Social Judgment *Motivation and Emotion* **22,** 187-209.

Dolan, R. J. (2002) Emotion, cognition, and behavior *Science* **298,** 1191-1194.

Dolan, R. J., Morris, J. S. and de Gelder, B. (2001) Crossmodal binding of fear in voice and face *Proceedings of the National Academy of Sciences of the United States of America.* **98,** 10006-10.

Doux, J. L. (1998) *The Emotional Brain,* Weidenfeld & Nicolson, Phoenix, Arizona.

DSM-IV (1994) *Diagnostic and statistical manual of mental disorder,* Washington DC.

Eisenberg, N. and Strayer, J. (Eds.) (1987) *Empathy and its development,* Cambridge University Press, Cambridge.

Ekman, P. (1999) Basic emotions.In *Handbook of Cognition and Emotion* (Eds).Dalgleish, T. and Powers, M. Wiley, New York.

Emery, N. J. (2000) The eyes have it: the neuroethology, function and evolution of social gaze *Neuroscience & Biobehavioral Reviews.* **24,** 581-604.

Eslinger, P. J. (1998) Neurological and neuropsychological basis of empathy *European Neurology* **39,** 193-199.

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarrier, N., Griffiths, P. D. and Woodruff, P. W. (2001) Investigating the functional anatomy of empathy and forgiveness *NeuroReport* **12,** 2433-8.

Fenigstein, A. and Abrams, D. (1993) Self-attention and the egocentric assumption of shared perspectives *Journal of Experimental Social Psychology* **29,** 287-303.

Fine, C., Lumsden, J. and Blair, R. J. R. (2001) Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain* **124,** 287-298.

Fisher, N. (2002) *The relationship of theory of mind to language and executive function in children with autism and children with moderate learning difficulties.* University of London

Fletcher, P., Happe, F., Frirth, U., Baker, S., Dolan, R., Frackowiak, R. and Frith, C. (1995) Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension *Cognition* **57,** 109-128.

Flint, A., Black, S. E., Campbell-TAylor, I., Gailey, G. F. and Levinton, C. (1993) Abnormal speech Articulation, psychomotor retardation, and subcortical dysfunction in major depression *Journal of Psychiatric Research* **27,** 309-319.

Frith, C. D. and Corcoran, R. (1996) Exploring theory of mind in people with schizophrenia *Psychological Medicine* **26,** 521- 530.

Frith, U. (2001) Mind blindness and the brain in autism *Neuron* **32,** 969-979.

Frith, U. (2003) *Autism: explaining the enigma,* Blackwell Publishing.

Frith, U. and Frith, C. D. (2003) Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society; Biological Sciences* **2003.**

Frith, U. and Happe, F. (1999) Theory of mind and self-consciousness: What is it like to be autistic? *Mind & Language* **14,** 1-22.

Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U. and Frith, C. D. (2000) Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks *Neuropsychologia.* **38,** 11-21.

Gallagher, H. L., Jack, A. I., Roepstorff, A. and Frith, C. D. (2002) Imaging the intentional stance in a competitive game *Neuroimage.* **16,** 814-21.

Gallese, V. (2003) The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* **36,** 171-180.

Gallese, V. and Goldman, A. (1998) Mirror neurones and the simulation theory of mind-reading *Trends in Cognitive Sciences* **2,** 493-501.

Gallup, G., Anderson, J. and Platek, S. (2003) Self-awareness, social intelligence and schizophrenia.In *The self in neuroscience and psychiatry.* (Ed)David, A. S. Cambridge University Press, Cambridge, pp. 147-165.

Gauthier, I., Tarr, M., Anderson, A., Skudlarski, P. and Gore, J. (1999) Activation of the middle fusiform 'face area' increase with expertise in recognizing novel objects. *Nature Neuroscience* **2,** 568-563.

George, N., Driver, J. and Dolan, R. J. (2001) Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing *Neuroimage.* **13,** 1102-12.

Gepner, B., Deruelle, C. and Grynfeltt, S. (2001) Motion and emotion: a novel approach to the study of face processing by young autistic children *Journal of Autism & Developmental Disorders* **31,** 37-45.

Gillberg, C. (1992) The Emmanuel Miller memorial lecture 1991 -
Autism and autistic-like conditions:
subclasses among disorders of empathy *Journal of Child Psychology and Psychiatry and Allied Disciplines* **33,** 813-842.

Gillott, A., Furniss, F. and Walter, A. (2001) Anxiety in high-functioning children with autism *Autism* **5,** 277-286.

Godfrey, H. P. and Knight, R. G. (1984) The validity of actometer and speech activity measures in the assessment of depressed patients *British Journal of Psychiatry* **145,** 159-163.

Goldie, P. (1999) How we think of others' emotions *Mind and Language* **14,** 394-423.

Goleman (1995) *Emotional Intelligence: Why it can matter more than IQ?,* Bantam Books, New York.

Gopnik, A. and Goldman, A. (1993) Knowing our minds *Behavioral and Brain Sciences* **16,** 1-28.

Grattan, L. M., Bloomer, R.H., Archambault, F.X., & Eslinger, P. (1994) Cognitive flexibility and empathy after frontal lobe lesion *Neuropsychiatry, Neuropsyhcology and Behavioural Neurology.* **7,** 251-259.

Grezes, J. and Decety, J. (2001) Functional anatomy of execution, mental simualtion, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping* **12,** 1-19.

Griffin, R. and Baron-Cohen, S. (2002) The intentional stance: developmental and neurocognitive perspectives.In *Daniel Dennett* (Eds).Brook, A. and Ross, D. Cambridge University Press, Cambridge.

Griffith, E. M., Pennington, B. F., Wehner, E. and Rogers, S. J. (1999) Executive functions in young children with autism *Child Development* **70,** 817-832.

Gur, R., Schroeder, L., Turner, T., McGrath, C., Chan, R., Turetsky, B., Alsop, D., Maldjian, J. and Gur, R. (2002) Brain Activation during facial emotion processing *Neuroimage* **16,** 651-662.

Guttman, H. and Laporte, L. (2002) Alexithymia, empathy, and psychological symptoms in a family context *Comprehensive Psychiatry* **43,** 448-455.

Gygax, P., Oakhill, J., Garnham and A, M. (2003) The representation of characters' emotional responses: Do readers infer specific emotions? *Cognition and Emotion* **17,** 413-428.

Haddon, M. (2003) *The curious incident of the dog in the night-time,* Jonathan Cape, London.

Hadjikhani, N. and DeGelder, B. (2003) Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology* **13,** 2201-2205.

Hadwin, J., Baron-Cohen, S., Howlin, Patricia, & Hill, K., (1996) Can we teach children with autism to understand emotions, belief or pretense. *Development and Psychopathology* **8,** 345-365.

Hall, G., Szechtman, H. and Nahmias, C. (2003) Enhanced salience and emotion recognition in autism: A PET study *American Journal of Psychiatry* **160,** 14391441.

Halligan, P. and David, A. S. (2001) Cognitive neuropsychiatry: towards a scientific psychopathology *Nature Neuroscience Reviews* **2,** 209-215.

Hamman, S., Ely, T., Hoffman, J. and Kilts, C. (2002) Activation of human amygdala in positive and negative emotion *Psychological Science* **13,** 135-141.

Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C. and Buchel, C. (2003) The human action recognition system and its relationship to Broca's area: an fMRI study *Neuroimage* **19,** 637-644.

Happe, F. (1994a) An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped and normal children and adults. *Journal of Autism & Developmental Disorders* **24**, 129-154.

Happe, F. (1994b) Weschler IQ profile and theory of mind in autism: a research note *Journal of Child Psychology & Psychiatry & Allied Disciplines* **35**, 1461-1471.

Happe, F. (1999) Autism: cognitive deficit or cognitive style *Trends in Cognitive Sciences* **3**, 216-222.

Happe, F., Ehlers, S., Fletcher, P., Frith, U., Johannson, M., Gillberg, C., Dolan, R., Frackowiak, R. and Frith, C. (1996) 'Theory of mind' in the brain. Evidence from a PET scan of Asperger syndrome *NeuroReport.* **8**, 197-201.

Hariri, A., Tessitore, A., Mattay, V., Fera, F. and Weinberger, D. (2002a) The amygdala response to emotional stimuli: a comparison of faces and scenes *Neuroimage* **17**, 317-323.

Hariri, A., Tessitore, A., Mattay, V., Fera, F. and Weinberger, D. (2002b) Human amygdala responses to fearful eyes *Neuroimage* **17**, 214-222.

Harris, P. (1989) *Children and emotion - The development of psychological understanding,* Blackwell, Oxford.

Hart, S. D., Hare, R. D. and Harpur, T. J. (1992) The Psychopathy Checklist--Revised (PCL-R): An overview for researchers and clinicians.In *Advances in psychological assessment*, Vol. 8 (Ed)Rosen, J., McReynolds, P.

Haxby, J. V., Hoffman, E. A. and Gobbini, M. I. (2002) Human neural systems for face recognition and social communication *Biological Psychiatry* **51,** 59-67.

Heaton, P., Hermelin, B. and Pring, L. (1999) Can children with autistic spectrum disorders perceive affect in music? An experimental investigation *Psychological Medicine.* **29,** 1405-10.

Heavey, L., Phillips, W., Baron-Cohen, S. and Rutter, M. (2000) The Awkward Moments Test: a naturalistic measure of social understanding in autism *Journal of Autism & Developmental Disorders.* **30,** 225-36.

Heider, F. and Simmel, M. (1944) An experimental study of apparent behaviour *American Journal of Psychology* **57,** 243-259.

Heiser, M., Iacobobi, M., Maeda, F., Maruc, J. and Mazziotta, J. (2003) The essential role of Broca's area in imitation *European Journal of Neuroscience* **17,** 1123-1128.

Henderson, D. (1996) Simulation theory vs theory theory - A difference without a difference in explanation *Southern Journal of Philosophy, Spindel Conference Supplement* **34,** 65-94.

Hobson, P. (1991) Methodological issues for experiments on autistic individuals' perception and understanding of emotion. *Journal of Child Psychology & Psychiatry & Allied Disciplines* **32,** 1135-1158.

Hobson, R. P. and Lee, A. (1999) Imitation and identification in autism *Journal of Child Psychology & Psychiatry & Allied Disciplines* **40,** 649-659.

Hobson, R. P., Ouston, J. and Lee, A. (1988) Emotion recognition in autism: Co-ordinating faces and voices *Psychological Medicine* **18,** 911-923.

Hoffman, E. and Haxby, J. (2000) Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience* **3,** 80 - 84.

Hoffman, M. L. (2000) *Empathy and moral development: Implications for caring and justice,* Cambridge University Press, New York.

Hogan, R. (1969) Development of an empathy scale *Journal of Consulting and Clinical Psychology* **33,** 307-316.

Hornak, J., Rolls, E.T. & Wade, D. (1996) Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage *Neuropsychologia* **34,** 247-261.

Houston, D. (1990) Empathy and the self: Cognitive and emotional influences on the evaluation of negative affect in others *Journal of Personality and Social Psychology* **59,** 859-868.

Howard, M. A., Cowell, P., Boucher, J., Broks, P., Mayes, A., Farrant, A. and Roberts, N. (2000) Convergent neuroanatomical and behaviour evidence of an amygdala hypothesis of autism. *NeuroReport.* **11,** 2931-2935.

Hughes, C., Russell, J. A. and Robbins, T. (1994) Evidence for executve dysfunction in autism *Neuropsychologia* **32,** 477-492.

Humphreys, G., Donnelly, N. and Riddoch, J. (1993) Expression is computed seperately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence. *Neuropsychologia* **31,** 173-181.

Hunter, E., Phillips, M., Chalder, T., Sierra, M. and David, A. S. (2003) Depersonalisation Disorder: A cognitive-behavioural conceptualisation *Behavior Research and Therapy* **41,** 1451-1467.

Hutcheson, G. and Sofroniou, N. (1999) *The multivariate social scientist,* Sage Publications Ltd, London.

ICD-10 (Tenth Edition) (Ed, Organisation, W. H.).

Ickes, W. (1997) *Empathic accuracy,* Guilford Press, New York.

Jansari, A., Tranel, D. and Adolphs, R. (2000) A valence-specific lateral bias for discriminating emotional facial expressions in free field *Cognition & Emotion* **14.**

Johnson, J., Smither, R. and Cheek, J. (1983) The structure of empathy *Journal of Personality and Social Psychology* **45,** 1299-1312.

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E. and Prigatano, G. P. (2002) Neural correlates of self-reflection *Brain.* **125,** 1808-14.

Joliffe, T. and Baron-Cohen, S. (1999) The strange stories test: a replication with high-functioning adults with autism or asperger syndrome. *Journal of Autism & Developmental Disorders* **29,** 395-404.

Jolliffe, T. and Baron-Cohen, S. (2001) A test of central coherence theory: Can adults with high-functioning autism or Asperger Syndrome integrate fragments of an object? *Cognitive Neuropsychiatry* **6,** 193-216.

Joseph, R. M. and Tanaka, J. (2003) Holistic and part-based face recognition in children with autism *Journal of Child Psychology & Psychiatry & Allied Disciplines* **44,** 529-542.

Kasari, C., Sigman, M., Baumgartner, P. and Stipek, D. (1993) Pride and mastery in children with autism *Journal of Child Psychology & Psychiatry & Allied Disciplines* **34,** 353-362.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., Fukuda, H., Kojima, S. and Nakamura, K. (1999) The human amygdala plays an important role in gaze monitoring. A PET study *Brain.* **122,** 779-83.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S. and Heatherton, T. F. (2002) Finding the self? An event-related fMRI study *Journal of Cognitive Neuroscience* **14,** 785-794.

Kemp, R., Chua, S., McKenna, P. and David, A. S. (1997) Reasoning and delusions *British Journal of Psychiatry* **170,** 398-405.

Kilner, J., Paulignan, Y. and Blakemore, S. J. (2003) An interference effect of observed biological movement on action. *Current Biology* **13,** 522-525.

Kilts, C., Egan, G., Gideo, D., Ely, T. and Hoffman, J. (2003) Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage* **18,** 156-168.

Kircher, T. T., Brammer, M., Bullmore, E., Simmons, A., Bartels, M. and David, A. S. (2002) The neural correlates of intentional and incidental self processing *Neuropsychologia.* **40,** 683-92.

Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., Simmons, A., Williams, S. C. R., Bartels, M. and David, A. S. (2000) Towards a functional neuroanatomy of self-processing: Effects of faces and words *Cognitive Brain Research* **10,** 133-144.

Kjaer, T. W., Nowak, M. and Lou, H. C. (2002) Reflective self-awareness and conscious states: PET evidence for a common midline parietofrontal core *Neuroimage.* **17,** 1080-6.

Klin, A. (2000) Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The Social Attribution Task *Journal of Child Psychology & Psychiatry & Allied Disciplines.* **41,** 831-46.

Klin, A., Jones, W., Schultz, W., Volkmar, F. and Cohen, D. (2002) Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism *Archives of General Psychiatry* **59,** 809-816.

Klin, A., Schultz, R. and Cohen, D. (2000) Theory of mind in action: developmental perspectives on social neuroscience.In *Understanding Other Minds: Perspectives from Developmental Neuroscience* (Ed)Cohen, D. Oxford University Press, Oxford.

Kucharska-Pietura, K., Phillips, M. L., Gernand, W. and David, A. S. (2003) Perception of emotions from faces and voices following unilateral brain damage *Neuropsychologia.* **41,** 1082-90.

LaBar, K., Crupain, M., Voyvodic, J. and McCarthy, G. (2003) Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex* **13,** 1023-1033.

Lambert, M. V., Senior, C., Fewtrell, W., Phillips, M. and David, A. S. (2001) Primary and secondary depersonalisation disorder: a psychometric study *Journal of Affective Disorders* **63,** 249-256.

Lane, R., Fink, G., Chua, M. and Dolan, R. (1997) Neural activation during selective attention to subjective emotional responses *Neuroreport* **8,** 3969-3972.

Lang, P., Greenwald, M., Bradley, M. and Hamm, A. (1993) Looking at pictures: affective, facial, visceral and behavioural reactions *Psychophysiology* **30,** 261-273.

Langdon, R. and Coltheart, M. (2001) Visual perspective-taking and schizotypy: evidence for a simulation-based account of mentalizing in normal adults. *Cognition* **82,** 1-26.

Langdon, R., Coltheart, M., Ward, P. and Catts, S. (2001) Visual and cognitive perspective-taking impairments in schizophrenia: a failure of allocentric simulation? *Cognitive Neuropsychiatry* **6,** 241-269.

Lawrence, E. J. and Peters, E. (2004 in press) Reasoning in believers in the paranormal. *Journal of Nervous & Mental Disease.*

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S. and David, A. S. (2004 in press) Measuring empathy - reliability and validity of the Empathy Quotient *Psychological Medicine* **34**.

Lawrence, E. J., Shaw, P., Dalton, J., Surguladze, S. and David, A. S. (2004) The online version of the revised Profile of Non-verbal Sensitivity. *In Preparation.*

Lawrence, K., Campbell, R., Swettenham, J., Terstegge, J., Akers, R., Coleman, M. and Skuse, D. (2003) Interpreting gaze in Turner Syndrome: impaired sensitivity to intention and emotion, but preservation of social cueing *Neuropsychologia.* **41,** 894-905.

Lee, A. and Hobson, P. (1994) I, you, me and autism: an experimental study *Journal of Autism & Developmental Disorders* **24,** 155-176.

Lidaka, T., Omori, M., Murata, T., Kosaka, H., Yonekura, Y., Okada, T. and Sadato, N. (2003) Neural interaction of the amygdala with the prefrontal and temporal cortices in the processing of facial expressions as revealed by fMRI *Journal of Cognitive Neuroscience* **18,** 1035-1047.

Loehlin, J. and Nichols, R. (1976) *Heredity, environment and personality: A study of 850 sets of twins.*, University of Texas Press, Austin, TX.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L. and Schopler, E. (1989) Autism Diagnostic Schedule: A standardised observation of communicative and social Behaviour *Journal of Autism & Developmental Disorders* **19,** 185 - 212.

Lord, C., Rutter, M. and Le Couteur, A. (1994) Autism Diagnostic Interview--Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders *Journal of Autism & Developmental Disorders* **24(5)**.

Loveland, K. A., Tunali-Kotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A. and Gibbs, M. C. (1997) Emotion recognition in autism: Verbal and nonverbal information *Development & Psychopathology* **9,** 579-593.

Luckett, T., Powell, S.D., Thornton, M.E. Schulz, J. (2002) Do children with autism who pass false belief tasks understand the mind as an active interpreter? *Journal of Autism & Developmental Disorders* **32,** 127-140.

Lutchmaya, S. and Baron-Cohen, S. (2002) Human sex differences in social and non-social looking preferences, at 12 months of age *Infant Behavior & Development* **25,** 319-325.

McCabe, K., Houser, D., Ryan, L., Smith, V. and Trouard, T. (2001) A functional imaging study of cooperation in two-person reciprocal exchange *Proceedings of the National Academy of Sciences of the United States of America.* **98,** 11832-5.

McGuire, P. K., Paulseu, E., Frackowiak, R. and Frith, C. D. (1996) Brain activation during stimulus independent thought *NeuroReport*. **7**, 2095-2099.

Mckenzie, K. S. and Hoyle, R. (2003) *The Self-Absorption Scale* University of Kentucky

Mearns, J. (2002) The autism/engineering link: a replication of Baron-Cohen et al (1997) *International Pediatrics* **17**, 60-61.

Medford, N., Baker, D., Hunter, H., Sierra, S., Lawrence, E. J., Phillips, M. and David, A. S. (2003) Chronic depersonalization following illicit drug use: Review of Forty Cases *Addiction* **98**, 1731-1736.

Mehrabian, A. and Epstein, N. (1972) A measure of emotional empathy *Journal of Personality* **40**, 525-543.

Mehrabian, A., Young, A. L. and Sato, S. (1988) Emotional empathy and associated individual differences *Current Psychology: Research & Reviews*, **7**, 221-240.

Mellor, C. (1988) Depersonalisation and self perception *British Journal of Psychiatry* **153**, 15-19.

Meltzoff, A. and Decety, J. (2003) What imitation tells us about social cognition: a rapproachment between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society; Biological Sciences* **358**, 491-500.

Meltzoff, A. N. and Gopnik, A. (1993) The role of imitation in understanding persons and developing a theory of mind.In *Understanding other minds: perspectives from autism* (Eds).Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D.

Morris, J. S., Scott, S. K. and Dolan, R. J. (1999) Saying it with feeling: neural responses to emotional vocalizations *Neuropsychologia.* **37,** 1155-63.

Nelson, H. E. (1982) *National Adult Reading Test (NART): Test Manual,* NFER: Wilson, Windsor.

Nezlek, J., Feist, G., Wilson, C. and Plesko, R. (2001) Day-to-day variability in empathy as a function of daily events and mood *Journal of Research in Personality* **35,** 401-423.

Noyes, R. and Kletti, R. (1977) Depersonalization in response to life-threatening danger *Comprehensive Psychiatry* **18,** 375-384.

Ogai, M., Matsumoto, H., Suzuki, K., Ozawa, F., Fukuda, R., Uchiyama, I., Suckling, J., Isoda, H., Mori, N. and Takei, N. (2003) fMRI study of recognition of facial expressions in high-functioning autistic patients *Neuroreport.* **14,** 559-63.

Ozonoff, S., Pennington, B. F. and Rogers, S. J. (1990) Are there emotion perception deficits in young autistic children? *Journal of Child Psychology & Psychiatry & Allied Disciplines* **31,** 343-361.

Ozonoff, S., Rogers, S. J. and Pennington, B. F. (1991) Asperger's Syndrome: Evidence of an empirical distinction from high functioning autism *Journal of Child Psychology & Psychiatry & Allied Disciplines* **32,** 1107-1122.

Parkinson, B. and Manstead, A. (1993) Making sense of emotion in stories and social life *Cognition and Emotion* **7,** 295-323.

Pelphrey, K., Singerman, J., Allison, T. and McCarthy, G. (2003) Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia* **41,** 156-170.

Philips, M., Young, A., Scott, S., Calder, A., Andrew, C., Giampietro, V., Williams, S., Bullmore, E., Brammer, M., & Gray, J. (1998) Neural Reponses to facial and vocal expressions of fear and disgust *Proceedings for the Royal Society London* **265,** 1809-1817.

Phillips, M., Drevets, W., Rauch, S. and Lane, R. (2003) Neurobiology of emotion perception 1: The neural basis of normal emotion perception *Biological Psychiatry* **54,** 504-514.

Phillips, M., Young, A., Scott, S., Calder, A., Andrew, C., Giampietro, V., Williams, , S., B., E., , Brammer, M. and Gray, J. (1998) Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society London* **265,** 1809-1817.

Phillips, M. L., Medford, N., Senior, C., Bullmore, E. T., Suckling, J., Brammer, M. J., Andrew, C., Sierra, M., Williams, S. C. R. and David, A. S. (2001)

Depersonalization disorder: Thinking without feeling *Psychiatry Research: Neuroimaging* **108,** 145-160.

Pierce, K., Muller, R., Ambrose, J., Allen, G. and Courchesne, E. (2001) Face processing occurs outside the fusiform 'face area' in autism: evidence from functional fMRI. *Brain* **124,** 2059-2073.

Pissiota, A., Frans, O., Michelgard, A., Appel, L., Langstom, B., Flaten, M. and Fredrikson, M. (2003) Amygdala and anterior cingulate cortex activation during affective startle modulation: a PET study of fear *European Journal of Neuroscience* **18,** 1325-1331.

Plutchik, R. (1980) A general psychoevolutionary theory of emotion.In *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion*, Vol. Vol.1 (Eds).Plutchik, R. and Kellerman, H. Academic, New York, pp. 3-33.

Popper, K. (1963) *Conjectures and refutations,* Routledge and Kegan Paul, London.

Premack, D. and Woodruff, G. (1978) Does the chimpanzee have a 'theory of mind'? *Behavioural and Brain Sciences* **4,** 515-526.

Preston, S. D. (2002) *A Perception-Action model of cognitive empathy? A PET Investigation of imagining your own experience and someone else's* Perspectives on Imitation France

Preston, S. D. and deWaal, F. B. (2003) Empathy: its ultimate and proximate basis *Behavioral and Brain Sciences* **25,** 1-72.

Prkachin, G. (2003) The effects of orientation on the detection and identification of facial expressions of emotion *British Journal of Psychology* **94,** 45-62.

Ramnani, N. and Miall, C. (2004) A system in the human brain for predicting the actions of others *Nature Neuroscience* **7,** 85-90.

Richell, R., Mitchell, D., Newman, C., Leonard, A., Baron-Cohen, S. and Blair, J. (2003) Theory of mind and psychopathy: can psychopathic individuals read the 'language of the eyes'? *Neuropschologia* **41,** 523-526.

Ring, H. A., Baron-Cohen, S., Wheelwright, S., Williams, S. C., Brammer, M., Andrew, C. and Bullmore, E. T. (1999) Cerebral correlates of preserved cognitive skills in autism: a functional MRI study of embedded figures task performance *Brain* **122,** 1305-15.

Roeyers, H., Buysse, A., Ponnet, K. and Pichal, B. (2001) Advancing advanced mind reading tests; Empathic accuracy in adults with pervasive developmental disorder *Journal of Child Psychology & Psychiatry & Allied Disciplines* **42,** 271-278.

Rogers, C. R. (1992) The processes of therapy *Journal of Consulting & Clinical Psychology* **60,** 163-164.

Rogers, S. J., Hepburn, S., Stackhouse, T. and Wehner, E. (2003) Imitation performance in toddlers with autism and those with other developmental disorders *Journal of Child Psychology & Psychiatry & Allied Disciplines* **44,** 763-781.

Rosenthal, H., Hall, J. A., DiMatteo, M. R., Rogers, P. L. and Archer, D. (1979) *Sensitivity to non-verbal communication: The PONS test,* Johns Hopkins University Press, Baltimore.

Ross, E. D., Thompson, R. D. and Yenkosky, J. (1997) Lateralization of affective prosody in brain and the callosal intergration of hemispheric language functions *Brain and Language* **56,** 27-54.

Rowe, A. D., Bullock, P. R., Polkey, C., E., and Morris, R. G. (2001) 'Theory of Mind' impairments and their relationship to excutive functioning following frontal lobe excisions. *Brain* **124,** 600-616.

Ruby, P. and Decety, J. (2001) Effect of subjective perspective taking during simulation of action: a PET investigation of agency *Nature Neuroscience* **4,** 546-50.

Ruby, P. and Decety, P. (2003) What you believe versus what you think they believe: a neuroimaging study of conceptual perspective-taking. *European Journal of Neuroscience* **17,** 2475-2480.

Russell, T., Rubia, K., Bullmore, E., Soni, W., Suckling, J., Brammer, M., Simmons, A., Williams, E. and Sharma, T. (2000) Exploring the social brain in schizophrenia: Left prefrontal underactivation during mental state attribution *American Journal of Psychiatry* **157,** 2040-2042.

Rutherford, M., Baron-Cohen, S. and Wheelwright, S. (2002) Reading the mind in the voice: a study with normal adults and adults with asperger syndrome and high functioning autism *Journal of Autism & Developmental Disorders* **32,** 189-194.

Salmond, C., Haan, M., Friston, K., Gadian, D. and Vargha-khadem, F. (2003) Investigating individual differences in brain abnormalities in autism *Philosophical Transactions of the Royal Society; Biological Sciences* **358,** 405-413.

Savoy, R. (2001) History and future directions of human brain mapping and functional neuroimaging. *Acta Psychologica* **107,** 9 - 42.

Scherer, K. (1995) Expression of emotion in voice and music *Journal of Voice* **9,** 235-248.

Scott, F., Baron-Cohen, S., Leslie, A. and Roth, D. (1999) 'If pigs could fly': A test of counterfactual reasoning and pretence in children with autism *British Journal of Developmental Psychology* **17,** 349-362.

Senior, C., Hunter, E., Lambert, M. V., Medford, N. C., Sierra, M., Phillips, M. L. and David, A. S. (2001) Depersonalisation *Psychologist* **14(3),** 128-132.

Serra, M., Loth, F., Van Geert, P., Hurkens, E. and Minderaa, R. (2002) Theory of mind in children with 'lesser variants' of autism: a longitudinal study *Journal of Child Psychology & Psychiatry & Allied Disciplines* **43,** 885-900.

Shallice, T. (1989) *From neuropsychology to mental structure,* Cambridge University Press, New York.

Shallice, T. (2003) Functional imaging and neuropsychology findings: how can they be linked? *Neuroimage* **20,** S146-S154.

Shalom, B., Mostofsky, R., Hazlett, R., Goldberg, M., McLeod, D. and Hoehn-Saric, R. (2003) Intact galvanic skin responses and impaired self-reports in response to emotional pictures in high-functioning autism *Annals of the New York Academy of Sciences* **985,** 501-504.

Shamay, S. G., Tomer, R., Berger, B. and Aharon-Peretz, J. (2003) Characterisation of empathy deficits following prefrotnal brain damage: The role of the right ventromeidal prefrontal cortex. *Journal of Cognitive Neuroscience* **15,** 324-337.

Shamay-Tsoory, S. G., Tomer, R., Yaniv, S. and Aharon-Peretz, J. (2002) Empathy deficits in Asperger Syndrome: A cognitive profile *Neurocase* **8,** 245-252.

Shaw, P., Lawrence, E., Baron-Cohen, S. and David, A. S. (2003a) Role of the amygdala in social sensitivity *Annals of the New York Academy of Sciences* **985,** 508-510.

Shaw, P., Lawrence, E., Radbourne, C., Bramham, J. and David, A. S. (2004) The impact of early and late damage to the human amygdala on 'theory of mind' reasoning. *Brain* ~~(submitted)~~. in press .

Shriberg, L., Paul, R., McSweeney, J., Klin, A., Cohen, D. and Volkmar, F. (2001) Speech and prosody characteristics of adolescents and adults with high-functioning autism and asperger syndrome. *Journal of Speech, Language and Hearing Research* **44,** 1097-1115.

Siegman, A. W. and Boyle, S. (1993) Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology* **102,** 430-437.

Sierra, M. and Berrios, G. E. (2001) The phenomenological stability of depersonalization: Comparing the old with the new *Journal of Nervous & Mental Disease* **189,** 629-636.

Sierra, M., Senior, C., Dalton, J., McDonough, M., Bond, A., Phillips, M. L., O'Dwyer, A. M. and David, A. S. (2002) Autonomic response in depersonalization disorder *Archives of General Psychiatry* **59,** 833-838.

Sigman, M., Kasari, C., Kwon, J. and Yirmiya, N. (1992) Responses to the negative emotions of others by autistic, mentally retarded and normal children. *Child Development* **63,** 796-807.

Silvia, P. J. and Abele, A. E. (2002) Can positive affect induce self-focused attention? Methodological and measurement issues *Cognition & Emotion* **16,** 845-853.

Simeon, D., Gross, S., Guralnik, O., Stein, D. J. and et al. (1997) Feeling unreal: 30 cases of DSM-III-R depersonalization disorder *American Journal of Psychiatry* **154,** 1107-1113.

Simeon, D., Guralnik, O., Gross, S., Stein, D. J., Schmeidler, J. and Hollander, E. (1998) The detection and measurement of depersonalization disorder *Journal of Nervous & Mental Disease* **186,** 536-542.

Sonnby-Borgstrom, M., Jonsson, P. and Svensson, P. (2003) Emotional empathy as related to mimicry reactions at different levels of information processing *Journal of Nonverbal Behaviour* **27**, 3-23.

Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacjer, P., Federlein, J., Buttner, T. and Przuntek, H. (1999) Knowing no fear *Proceedings for the Royal Society London* **266**, 2451-2456.

Stone, V. (1999) The role of the frontal lobes and amygdala in theory of mind.In *Understanding Other Minds - Perspectives from Developmental Cognitive Neuroscience* (Eds).Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D.

Stone, V., Baron-Cohen, S. and Knight, R. (1998) Frontal lobe contributions to theory of mind *Journal of Cognitive Neuroscience* **10**, 640-656.

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J. and Young, A. (2003) Acquired theory of mind impairments in individuals with bilateral amygdala lesions *Neuropsychologia* **41**, 209-220.

Stone, W., Ousley, O. and Littleford, C. (1997) Motor imitation in young children with autism: what's the object? *Journal of Abnormal Psychology* **25**, 475-485.

Surguladze, S., Brammer, M., Young, A. W., Andrew, C., Travis, M., Williams, S. and Phillips, M. (2003) A preferential increase in the extrastraite response to signals of danger *Neuroimage* **19**, 1317-1328.

Swallow, K., Braver, T., Snyder, A., Speer, N. and Zacks, J. (2003) Reliability of functional localisation using fMRI *Neuroimage* **20,** 1561-1577.

Szatmari, P., Bartolucci, G. and Bremner, R. (1989) Asperger's Syndrome and autism: a comparison of early history and outcome *Developmental Medicine and Child Neurology* **31,** 709-720.

Tabachnick, B. and Fiddell, L. (1989) *Using multivariate statistics,* Harper Collins, New York.

Tager-Flusberg, H. (1992) Autistic children's talk about psychological states: deficits in the early acquisition of theory of mind *Child Development* **63,** 161-172.

Tager-Flusberg, H. and Sullivan, K. (2000) A componential view of theory of mind: evidence from Williams Syndrome *Cognition* **76,** 59-89.

Talairach, J. and Tournoux, P. (1988) *Co-planar stereotaxic atlas of the human brain.,* Thieme Verlag, Stuttgart.

Taylor, S., Phan, K., Decker, L. and Liberzon, I. (2003) Subjective ratings of emotionally salient stimuli modulates neural activity *Neuroimage* **18,** 650-659.

Teasdale, J. D. and Fogarty, S. J. (1979) Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory *Journal of Abnormal Psychology* **88(3),** 248-257.

Teasdale, J. D., Fogarty, S. J. and Williams, M. (1980) Speech rate as a measure of short term variation in depression *British Journal of Social and Clinical Psychology* **19,** 271-278.

Teunisse, J. and de Gelder, B. (2003) Face processing in adolescents with autistic disorder: The inversion and composite effects *Brain and Cognition* **52,** 285-294.

Toichi, M., Kamio, Y., Okada, T., Sakihama, M., Youngstrom, E. A., Findling, R. L. and Yamamoto, K. (2002) A lack of self-consciousness in autism *American Journal of Psychiatry.* **159,** 1422-4.

Toomey, R., Schuldberg, D., Corrigan, P. and Green, M. F. (2002) Nonverbal social perception and symptomatology in schizophrenia *Schizophrenia Research.* **53,** 83-91.

Vaina, L., Soloman, J., Chowdhury, S., Sinha, P. and Belliveau, J. (2001) Functional neuroanatomy of biological motion perception in humans *Proceedings for the National Aacdemy of Science, USA* **98,** 11656-11661.

Van der Mark, I., Van Ijzendoorn, M. and Bakermans-Kranenburg, M. (2002) Development of empathy in girls during the second year of life: Associations with parenting, attachment and temperament *Social Development* **11,** 451-468.

Vogeley, K., Bussfeld, S. and Newen, A. (2001) Mind Reading: Neural mechanisms of theory of mind and self-perspective *Neuroimage* **14,** 170-181.

Vogeley, K. and Fink, G. R. (2003) Neural correlates of first person perspective *Trends in Cognitive Sciences* **7,** 38-42.

Vrana, S. (1988) The startle response probe: A new measure of emotion *Journal of Abnormal Psychology* **97,** 487-491.

Weiss, U., Salloum, J. and Scneider, F. (1999) Correspondance of emotional self-rating with facial expression *Psychiatry Research* **86,** 175-184.

Whalen, P., Rauch, S., Etcoff, N., McInerney, S., Lee, M., & Jenike, M. (1998) Masked presenations of emotional facial expressions modulate amygdala activity without explicit knowledge *Journal of Neuroscience* **18,** 411-418.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V. and Rizzolatti, G. (2003a) Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust *Neurone* **40,** 655-664.

Wicker, B., Perrett, D., I., Baron-Cohen, S. and Decety, J. (2003b) Being the target of anothers emotion: a PET study *Neuropsychologica* **41,** 139-146.

Williams, J., Whiten, A., Suddendorf, T. and Perrett, D. (2001) Imitation, mirror neurones and autism *Neuroscience and Biobehavioural Review* **25,** 287-295.

Wimmer, H. and Perner, J. (1983) Beliefs abut beliefs: representation and constraining function of wrong belief's in young children's understanding of deception. *Cognition* **13,** 103-28.

Yirmiya, N., Sigman, M., Kasari, C. and Mundy, P. (1992) Empathy and cognition in high-functioning children with autism *Child Development* **63,** 150-160.

Zahavi, D. and Parnas, J. (2003) Conceptual problems in infantile autism research *Journal of Consciousness Studies* **10,** 53-71.

Zald, D. (2003) The human amygdala and the emotional evaluation of sensory stimuli *Brain Research Reviews* **41,** 88-123.

Appendices

Appendix 1:The Empathy Quotient – (Baron-Cohen and Wheelwright, 2004 in press).

How to fill out the questionnaire

Below are a list of statements. Please read each statement very carefully and rate how strongly you agree or disagree with it by circling your answer. There are no right or wrong answers, or trick questions.

IN ORDER FOR THE SCALE TO BE VALID, YOU MUST ANSWER EVERY QUESTION.

Examples

E1. I would be very upset if I couldn't listen to music every day.	strongly agree	slightly agree	slightly disagree	strongly disagree
E2. I prefer to speak to my friends on the phone rather than write letters to them.	strongly agree	slightly agree	slightly disagree	strongly disagree
E3. I have no desire to travel to different parts of the world.	strongly agree	slightly agree	slightly disagree	strongly disagree
E4. I prefer to read than to dance.	strongly agree	slightly agree	slightly disagree	strongly disagree

PLEASE TURN OVER

1. I can easily tell if someone else wants to enter a conversation.	strongly agree	slightly agree	slightly disagree	strongly disagree
2. I prefer animals to humans.	strongly agree	slightly agree	slightly disagree	strongly disagree
3. I try to keep up with the current trends and fashions.	strongly agree	slightly agree	slightly disagree	strongly disagree
4. I find it difficult to explain to others things that I understand easily, when they don't understand it first time.	strongly agree	slightly agree	slightly disagree	strongly disagree
5. I dream most nights.	strongly agree	slightly agree	slightly disagree	strongly disagree
6. I really enjoy caring for other people.	strongly agree	slightly agree	slightly disagree	strongly disagree
7. I try to solve my own problems rather than discussing them with others.	strongly agree	slightly agree	slightly disagree	strongly disagree
8. I find it hard to know what to do in a social situation.	strongly agree	slightly agree	slightly disagree	strongly disagree
9. I am at my best first thing in the morning.	strongly agree	slightly agree	slightly disagree	strongly disagree
10. People often tell me that I went too far in driving my point home in a discussion.	strongly agree	slightly agree	slightly disagree	strongly disagree
11. It doesn't bother me too much if I am late meeting a friend.	strongly agree	slightly agree	slightly disagree	strongly disagree
12. Friendships and relationships are just too difficult, so I tend not to bother with them.	strongly agree	slightly agree	slightly disagree	strongly disagree

13. I would never break a law, no matter how minor.	strongly agree	slightly agree	slightly disagree	strongly disagree
14. I often find it difficult to judge if something is rude or polite.	strongly agree	slightly agree	slightly disagree	strongly disagree
15. In a conversation, I tend to focus on my own thoughts rather than on what my listener might be thinking.	strongly agree	slightly agree	slightly disagree	strongly disagree
16. I prefer practical jokes to verbal humour.	strongly agree	slightly agree	slightly disagree	strongly disagree
17. I live life for today rather than the future.	strongly agree	slightly agree	slightly disagree	strongly disagree
18. When I was a child, I enjoyed cutting up worms to see what would happen.	strongly agree	slightly agree	slightly disagree	strongly disagree
19. I can pick up quickly if someone says one thing but means another.	strongly agree	slightly agree	slightly disagree	strongly disagree
20. I tend to have very strong opinions about morality.	strongly agree	slightly agree	slightly disagree	strongly disagree
21. It is hard for me to see why some things upset people so much.	strongly agree	slightly agree	slightly disagree	strongly disagree
22. I find it easy to put myself in somebody else's shoes.	strongly agree	slightly agree	slightly disagree	strongly disagree
23. I think that good manners are the most important thing a parent can teach their child.	strongly agree	slightly agree	slightly disagree	strongly disagree
24. I like to do things on the spur of the moment.	strongly agree	slightly agree	slightly disagree	strongly disagree

PLEASE TURN OVER

25. I am good at predicting how someone will feel.	strongly agree	slightly agree	slightly disagree	strongly disagree
26. I am quick to spot when someone in a group is feeling awkward or uncomfortable.	strongly agree	slightly agree	slightly disagree	strongly disagree
27. If I say something that someone else is offended by, I think that that's their problem, not mine.	strongly agree	slightly agree	slightly disagree	strongly disagree
28. If anyone asked me if I liked their haircut, I would reply truthfully, even if I didn't like it.	strongly agree	slightly agree	slightly disagree	strongly disagree
29. I can't always see why someone should have felt offended by a remark.	strongly agree	slightly agree	slightly disagree	strongly disagree
30. People often tell me that I am very unpredictable.	strongly agree	slightly agree	slightly disagree	strongly disagree
31. I enjoy being the centre of attention at any social gathering.	strongly agree	slightly agree	slightly disagree	strongly disagree
32. Seeing people cry doesn't really upset me.	strongly agree	slightly agree	slightly disagree	strongly disagree
33. I enjoy having discussions about politics.	strongly agree	slightly agree	slightly disagree	strongly disagree
34. I am very blunt, which some people take to be rudeness, even though this is unintentional.	strongly agree	slightly agree	slightly disagree	strongly disagree
35. I don't tend to find social situations confusing.	strongly agree	slightly agree	slightly disagree	strongly disagree
36. Other people tell me I am good at understanding how they are feeling and what they are thinking.	strongly agree	slightly agree	slightly disagree	strongly disagree

PLEASE TURN OVER

37. When I talk to people, I tend to talk about their experiences rather than my own.	strongly agree	slightly agree	slightly disagree	strongly disagree
38. It upsets me to see an animal in pain.	strongly agree	slightly agree	slightly disagree	strongly disagree
39. I am able to make decisions without being influenced by people's feelings.	strongly agree	slightly agree	slightly disagree	strongly disagree
40. I can't relax until I have done everything I had planned to do that day.	strongly agree	slightly agree	slightly disagree	strongly disagree
41. I can easily tell if someone else is interested or bored with what I am saying.	strongly agree	slightly agree	slightly disagree	strongly disagree
42. I get upset if I see people suffering on news programmes.	strongly agree	slightly agree	slightly disagree	strongly disagree
43. Friends usually talk to me about their problems as they say that I am very understanding.	strongly agree	slightly agree	slightly disagree	strongly disagree
44. I can sense if I am intruding, even if the other person doesn't tell me.	strongly agree	slightly agree	slightly disagree	strongly disagree
45. I often start new hobbies but quickly become bored with them and move on to something else.	strongly agree	slightly agree	slightly disagree	strongly disagree
46. People sometimes tell me that I have gone too far with teasing.	strongly agree	slightly agree	slightly disagree	strongly disagree
47. I would be too nervous to go on a big rollercoaster.	strongly agree	slightly agree	slightly disagree	strongly disagree
48. Other people often say that I am insensitive, though I don't always see why.	strongly agree	slightly agree	slightly disagree	strongly disagree

PLEASE TURN OVER

49. If I see a stranger in a group, I think that it is up to them to make an effort to join in.	strongly agree	slightly agree	slightly disagree	strongly disagree
50. I usually stay emotionally detached when watching a film.	strongly agree	slightly agree	slightly disagree	strongly disagree
51. I like to be very organised in day to day life and often make lists of the chores I have to do.	strongly agree	slightly agree	slightly disagree	strongly disagree
52. I can tune into how someone else feels rapidly and intuitively.	strongly agree	slightly agree	slightly disagree	strongly disagree
53. I don't like to take risks.	strongly agree	slightly agree	slightly disagree	strongly disagree
54. I can easily work out what another person might want to talk about.	strongly agree	slightly agree	slightly disagree	strongly disagree
55. I can tell if someone is masking their true emotion.	strongly agree	slightly agree	slightly disagree	strongly disagree
56. Before making a decision I always weigh up the pros and cons.	strongly agree	slightly agree	slightly disagree	strongly disagree
57. I don't consciously work out the rules of social situations.	strongly agree	slightly agree	slightly disagree	strongly disagree
58. I am good at predicting what someone will do.	strongly agree	slightly agree	slightly disagree	strongly disagree
59. I tend to get emotionally involved with a friend's problems.	strongly agree	slightly agree	slightly disagree	strongly disagree
60. I can usually appreciate the other person's viewpoint, even if I don't agree with it.	strongly agree	slightly agree	slightly disagree	strongly disagree

Thank you for filling this questionnaire in. © SBC/SJW Feb 1998

Appendix 2: 'Reading the Mind in the Eyes' Stimulus (Baron-Cohen *et al.*, 2001)

jealous panicked

arrogant hateful

Appendix 3: Diary Extracts Used in the Mental State Attribution Task

This is an extract from the diary of James who is 42 years old......

I was thinking today how the internet has changed my life in so many ways...I can't think of any other invention as of late that has as many uses as the internet. Ellie prefers to use the library and says the internet makes people less sociable. She just sees how long I spend on the computer – she doesn't realise it promotes communication. For me, the biggest attribute of the internet is to be able to access any and all types of information instantly. Since I do a lot of long distance walking, being able to locate specific routes is brilliant. The same goes for classic car things, as well. I belong to five classic car clubs, and I met all of those groups through the internet. Ellie thinks the internet isn't any quicker in the long run but it is if you know how to use it properly.

(extract adapted from opendiary.com, 2001)

1. James thinks that Ellie believes .....

a) the internet has made him sociable.
b) the internet has made him less sociable **X**[30]
c) the internet has made him unsociable.
d) the internet has made him more sociable.

2. Ellie believes the internet ....

a) is not faster than the library **X**
b) is slightly faster than the library.
c) is much faster than the library.
d) is faster than the library.

3. James uses the....

a) internet to find walking routes. **X**
b) internet to find running routes.
c) library to find running routes.
d) library to find walking routes.

4. James believes Ellie....

a) does know how to use the internet properly.
b) doesn't know how to use the internet at all.
c) does know how to use the internet competently.
d) doesn't know how to use the internet properly. **X**

[30] For ease of reference X denotes to 'preferred' answer according to the final set of control data.

This is an extract from the diary of Sally who is 24 years old....

Tomorrow is the funeral of my Nan, so as you can imagine I am really not looking forward to it. The hardest part about it is that when I see that coffin tomorrow I will know that this is the last time I will ever be that close to my Nan again. I hate the thought of that. Why did you have to die? I hate never being able to see you again. I hate that we will all be sitting in your house tomorrow and you won't be there. She loved that house. I just can't help how much I will miss her, I loved her so much. I know these are selfish thoughts, because I know how much she missed my Pop and she just wanted to be with him, so I should be happy for her. I can't believe this has happened. She was so young for her age, and fit and healthy.

(extract adapted opendiary.com, 2001)

1. Sally is....
a) calm even though her Nan has died.
b) angry because her Nan has died. **X**
c) furious because her Nan has died.
d) relaxed even though her Nan has died.

2. Sally is.....
a) very unhappy as she will miss her Nan. **X**
b) not unhappy even though she will miss her Nan.
c) inconsolable as she will miss her Nan.
d) consolable even though she will miss her Nan.

3. Sally will go to her.....
a) Pop's after the funeral.
b) mum's after the funeral.
c) Nan's after the funeral. **X**
d) brother's after the funeral.

4. Sally feels...........
a) ashamed because she isn't happy for her Nan.
b) unashamed even though she isn't happy for her Nan.
c) guiltless even though she isn't happy for her Nan.
d) guilty because she isn't happy for her Nan. **X**

This is an extract from the diary of Pauline who is 36 years old.......

I woke up this morning in the recovery room to a male nurse standing over me saying "Mrs. ---, did they tell you that you have a son?" I couldn't believe it – a boy as I'd hoped - I wanted to ask how he was, but I couldn't even get my eyes open, let alone my mouth. Eventually, I croaked out "Is he ok?" "Yes", the nurse said, "he was pretty big, too." Soon after that they took me to see him – he was in Intensive Care Unit as a precaution because he was born with a fever. He had "TONS" of long, black hair. This surprised us both since we were both born with moderate amounts of medium coloured hair. My baby boy was 8 pounds 3 ounces and 22 inches long. He looked just like my husband and I just wanted to hold him and not let go.

(extract adapted opendiary.com, 2001)

1. Pauline feels....
a) joyless despite having just given birth to a baby boy.
b) overjoyed having just given birth to a baby boy. **X**
c) touched having just given birth to a baby boy.
d) untouched despite having just given birth to a baby boy.

2. Pauline behaved in a ....
a) caring way towards her son.
b) loving way towards her son. **X**
c) loveless way towards her son.
d) uncaring way towards her son.

3. Pauline feels ....
a) fulfilled as she has a healthy baby. **X**
b) content as she has a healthy baby.
c) unfulfilled despite having a healthy baby.
d) discontent despite having a healthy baby.

4. The baby was born with ....
a) lots of medium coloured hair.
b) a small amount of medium coloured hair.
c) lots of dark coloured hair. **X**
d) a small amount of dark coloured hair.

This is an extract from the diary of Geoff who is 58 years old.......

I usually run at the sight of needles – but its too late for that now – there's no getting out of it. I can see lots of needles. The guys in front of me are being injected one at a time alternating each arm - left then right. "Oh crap - I'm next". The technician swabs my left arm with a cotton ball dipped in alcohol - that sickening smell of alcohol. I looked at him, the needle, then looked away. It seemed like an eternity before I felt the intense sting of this contraption inject its contents into my arm. He removed it from my arm and some residual fluid ran down. It was cold. The process continued all over again. Then two more times after that. By the time I was finished I felt like a pin cushion full of burning holes. I can remember turning pale and having to sit down.

(extract adapted opendiary.com, 2001)

1. Geoff had ....
a) two injections administered to each leg.
b) one injection administered to each leg.
c) two injections administered to each arm. **X**
d) one injection administered in each arm.

2. Geoff is.....
a) unhysterical before receiving an injection.
b) afraid before receiving an injection. **X**
c) hysterical before receiving an injection.
d) unafraid before receiving an injection.

3. Geoff finds the injection process ....
a) troubling from start to finish.
b) unbearable from start to finish. **X**
c) untroubling from start to finish.
d) bearable from start to finish.

4. Geoff is ....
a) unbothered when he realises he has to have an injection.
b) alarmed when he realises he has to have an injection.
c) bothered when he realises he has to have an injection. **X**
d) unalarmed when he realises he has to have an injection.

Appendix 4: Traits used in the Trait-Overlap Task

Adaptable
Alert
Ambitious
Anxious
Assertive
Capable
Cautious
Cheerful
Confident
Confused
Conscientious
Conventional

Curious
Defensive
Dependable
Dependent
Determined
Dissatisfied
Dominant
Easy Going
Efficient
Egotistical
Emotional
Energetic
Enthusiastic

Excitable
Fearful
Feminine
Friendly
Generous
Good-natured
Helpful
Highly-Strung
Honest
Humorous
Idealistic
Impulsive

Independent
Industrious
Inhibited
Insightful
Intelligent
Irresponsible
Irritable
Kind
Lazy
Logical
Loyal
Mature
Mischievous

Moody
Nervous
Obliging
Opinionated
Outgoing
Outspoken
Patient
Pleasant
Quiet
Realistic
Resourceful
Responsible

Self-centred
Sensitive
Serious
Shy
Sociable
Suspicious
Tense
Thoughtful
Timid
Warm
Withdrawn
Witty
Worrying

Appendix 5: First order ToM task (Rowe *et al.*, 2001)

It was Saturday morning and Cathy was going shopping. She drove into town quite early in the hope that she wouldn't have too much difficulty parking, but when she got there she was surprised to find that it was already busy with lots of shoppers around. It looked as though all the parking spaces were taken and she'd have to go to the outskirts of the town centre.

Meanwhile, another shopper, Mr. Watson, had just returned to his car. He had came to pick up some spectacles from the opticians'. Unsure of how long he would be, Mr. Watson had put sufficient change into his parking meter to allow the maximum two hour period for parking. He was pleased to find that it had only taken a few minutes to complete his business at the opticians' and he was back at his car within ten minutes.

Cathy had nearly given up searching for a parking space, but just then she saw Mr. Watson's car pulling out from the curb up ahead. Cathy was pleased; finding this space saved her quite a long walk. She parked the car, switched off the ignition, and got her purse out of her handbag, but she found she had no change for the parking meter. Traffic wardens often wandered up and down the main road and so Cathy got out of the car and, without looking at the meter, ran to the large newsagents' on the corner. As soon as Cathy had been served she rushed out with some coins in her hand and dashed back to the car.

Please answer the questions below as quickly and accurately as possible without referring back to the extract. We need you to answer all the questions. Please write your answers in the space provided.

1. Why was Cathy in such a hurry to put money in the meter?
2. Why did Cathy go into the newsagents?
3. Was there any time left on the parking meter?
4. What was the purpose of Mr.Watson's visit to town that morning?

Appendix 6: Second order ToM task (Rowe *et al.*, 2001)

Richard and Ann want to re-decorate their spare bedroom. They have several wallpaper samples which they have brought home from the DIY shop. After choosing one they decide to make a start the next weekend. Ann says she can collect the wallpaper on her way home from work next Friday evening.

On Friday morning Richard's boss unexpectedly tells him he can take the afternoon off. When he gets home he phones Ann at work. Unfortunately she is not available. One of Ann's colleagues tells him she is at a meeting and she is not sure when Ann will be back. Richard explains that he wants to tell her he's got the afternoon off and he's going to make a start on their re-decorating. He says, "Can you ask Ann to call me at home when she returns?"

By the time Ann gets back to her office it is after 5pm and all her colleagues have gone home. As she collects her things she sees a note on her desk saying, firstly, that her husband called to tell her he had the afternoon off and was beginning their redecorating work and, secondly, that she should phone him at home. It is late and Ann wants to get to the DIY shop before 6pm so she decides that, since she already has the message, she won't bother to phone Richard and she will just rush down to buy the wallpaper.

At home Richard has been working very hard. He hasn't heard from Ann and he wonders why her colleagues are so unreliable. He's nearly finished stripping off the old wallpaper when he hears Ann coming through the front door. He rushes downstairs and says "You're going to be surprised when I tell you what I've been doing this afternoon".

Please answer the questions below as quickly and accurately as possible without referring back to the extract. We need you to answer all the questions. Please write your answers in the space provided.

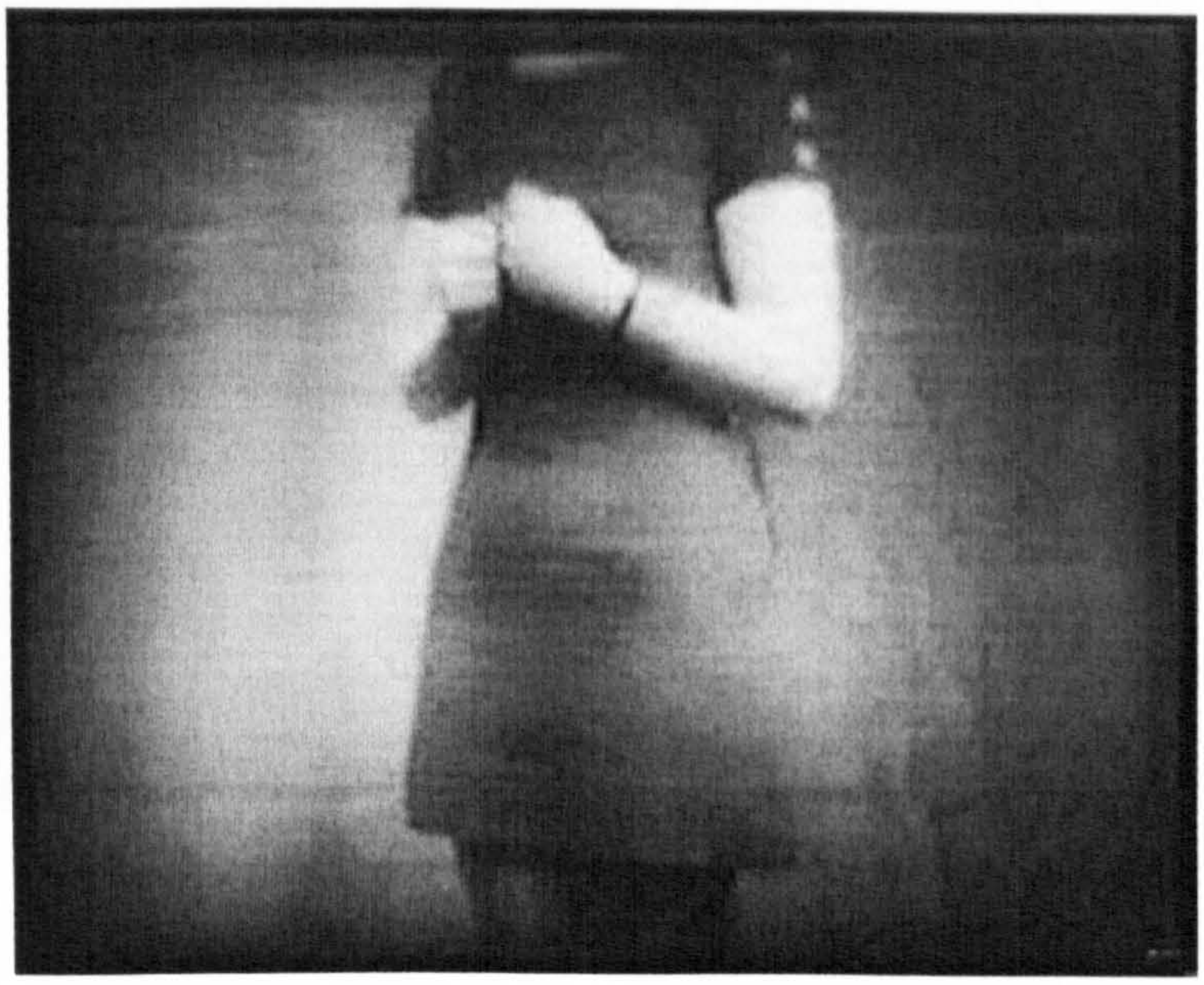
1. Why does Richard say that Ann is going to be surprised?
2. Why do you think Ann needed to hurry to get to the DIY shop?
3. Did Ann get Richard's message about having the afternoon off?
4. Which room were they re-decorating?

Appendix 7: Mental state labels used in PONS-r

Old Label	*New Label*
helping customer	explaining something
ordering food	ordering something
expressing gratitude	thankful
deep affection	deep affection
seducing someone	flirtatious
talking about ones wedding	very happy
leaving on a trip	saying goodbye
admiring nature	delighted
talking to a lost child	concerned
talking about the death of a friend	sombre
talking about ones divorce	serious discussion
returning a item to a shop	complaining
asking forgiveness	apologising
praying	pleading
criticising someone for being late	criticising
nagging a child	nagging
expressing strong dislike	exasperation
threatening someone	threatening someone
jealous anger	anger

Appendix 8: Stimuli used in PONS-r

Body - Complaining*



Face - Criticising

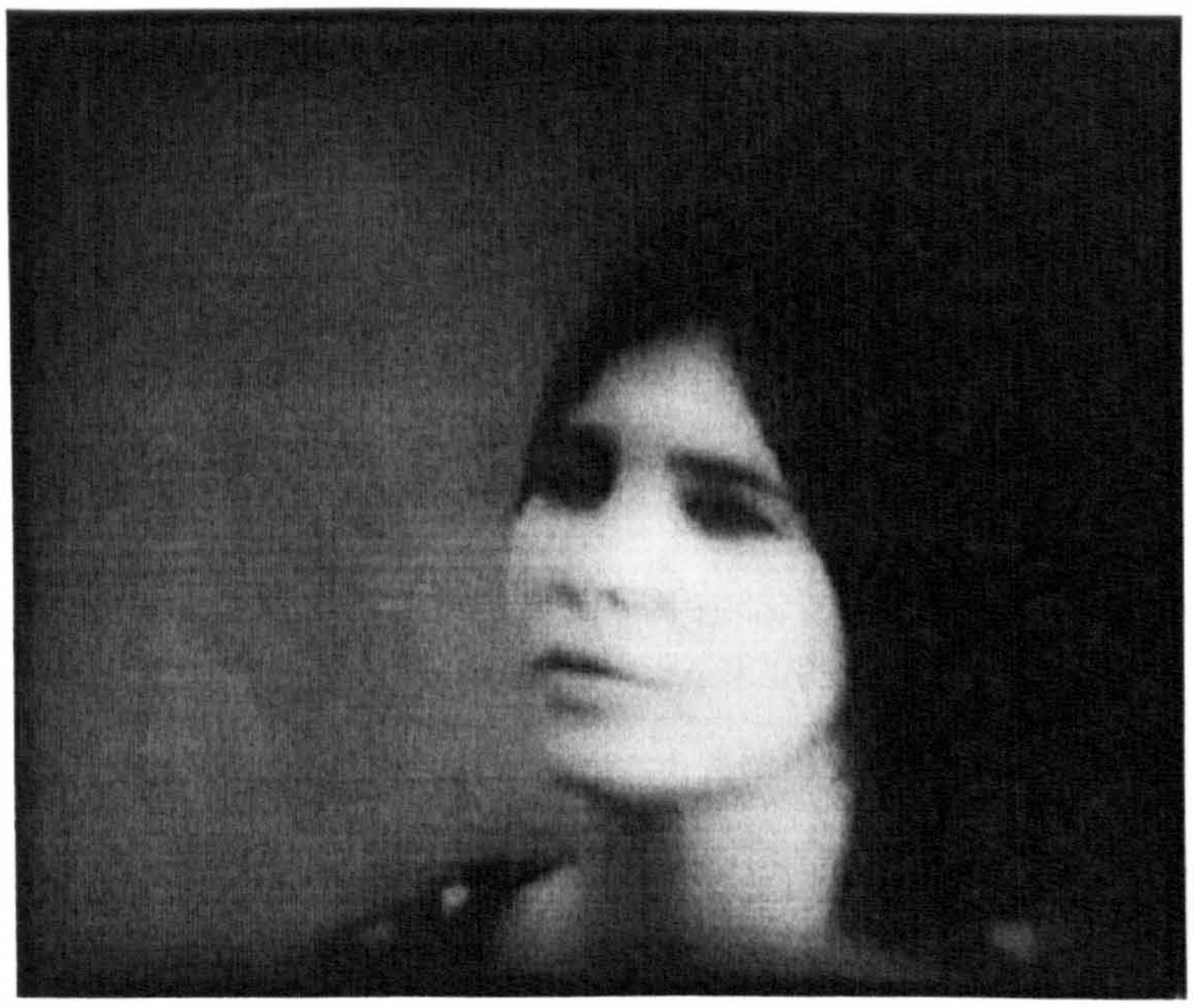


Figure - Anger

* Although the answers are given they may not appear congruent as the stimuli is meant to be viewed dynamically and develops over time